

Lab Assignment 4: Using APIs in Python

DS 6001: Practice and Application of Data Science

Instructions

Please answer the following questions as completely as possible using text, code, and the results of code as needed. Format your answers in a Jupyter notebook. To receive full credit, make sure you address every part of the problem, and make sure your document is formatted in a clean and professional way.

In this lab, you will work with the public API provided by genius.com, a website that calls itself "the world's biggest collection of song lyrics and musical knowledge." You will need to read the API documentation carefully, acquire an access key, and use it without sharing it to pull data from this API into Python. You will also practice using a library made specifically to wrap around `requests` to make calling from the Genius API easier.

Problem 0

Import the following libraries:

```
In [1]: import numpy as np
import pandas as pd
import requests
import json
import os
import dotenv
import sys
sys.tracebacklimit = 0 # turn off the error tracebacks
```

Problem 1

The Genius API documentation is here: <https://docs.genius.com/#/getting-started-h1>. Read through the documentation carefully. Although the Genius API is free and public, it still requires users an access key to use the API. In this case, Genius provides users with three codes: a client ID, a client ID secret, and a client access token. Use the documentation to find a way to obtain these codes for yourself. Write a paragraph that describes all of the steps you needed to take (but DO NOT list your access codes in this paragraph).

Some hints and cautions:

1. Before you can use the API, you will need a regular, free user account with Genius. Sign-up here: https://genius.com/signup_or_login
2. Genius's API is built to support third-party app development, not data scientists. The language is entirely geared toward app development. Under "Authentication" there are

instructions to third-party developers for guiding their own users in getting access keys. That's not relevant to getting access for yourself. It's not hard to get an access key, but the guidance here is not very clear. Be patient and read everything in the Authentication section carefully.

3. When you arrive at the page that allows you to register for API access keys, the language is still geared toward app development. You will be prompted to name your app and provide the URLs associated with the app. It doesn't much matter what you name your app, and I just used the Collab main page (<https://collab.its.virginia.edu/portal?containerLogin=true>) for the URLs.
4. When you get your codes, copy them in a text file. In problem 2 you will copy these codes over again to a `.env` file. [4 points]

Problem 2

Create a `.env` file for this project. Open it, copy your access codes into it, and save it. Then use Python code to load the environmental variables contained in the `.env` file, and create variables that contain each of the three codes. (You can print these variables to make sure it worked, but do not allow your access keys to display in your notebook). [4 points]

Problem 3

The root for all Genius APIs is <https://api.genius.com>. Find the endpoint for the Search API. (You will have to click the "Authorize with Genius" button in the upper-right corner if you haven't already done so). Use the `requests` library to issue a search for Bob Dylan. Genius's API is organized in a way that every individual artist has his or her own API endpoint. Display a portion of the JSON output that displays the API endpoint path for the data on Bob Dylan.

Hint: to authenticate, specify your access token (not your client ID or client secret) as the `access_token` parameter. You will have to dig around the JSON output to find the artist ID, but it is listed under `primary_artist` several branches down the JSON tree. [4 points]

Problem 4

Add `/songs` to the end of the the endpoint path you found in problem 3 and use this path to request the 20 most popular Bob Dylan songs. Organize these data in a `pandas` data frame. [4 points]

Problem 5

Install and import the `lyricsgenius` library in Python, which is a wrapper around `requests` that works specifically with the Genius API. Follow the guide on the GitHub repository for this library (<https://github.com/johnwmillr/LyricsGenius>) for instructions on using the library. Use the

lyricsgenius library to download and display the lyrics to "Tangled Up in Blue" by Bob Dylan. [4 points]