Project 7 — CIA Flag Images

CS 1410

Background

In this assignment, you will be downloading GIFs for flags from all the countries in the world from a public CIA website. The files appear in the following web folder:

https://www.cia.gov/library/publications/resources/the-world-factbook/graphics/flags/large/

The names of the files have the following form:

        xx-lgflag.gif

where **xx** stands for the 2-character country code. For example, the GIF for the United States is *us-lgflag.gif*. The 2-character codes are in the file **flags.txt** in the Canvas folder for this assignment.

Note: the web folder above is protected so you'll get a 404 error if you try to read the page directly. If you add the file name to the URL, it works fine. For example:

https://www.cia.gov/library/publications/resources/the-world-factbook/graphics/flags/large/us-lgflag.gif

Write three similar scripts that do the following:

- Download all flag files into a "flags" subdirectory
- Report the total number of bytes downloaded
- Report the execution time of the script (recorded via **time.perf_counter()**)

The three versions will use the following schemes for downloading:

1. Download *sequentially* (no concurrency) — name your file *cia_a.py*
2. Download concurrently using *processes* — file *cia_b.py*
3. Download concurrently using *threads* — file *cia_c.py*

Use the **requests** module discussed in class to make the HTTP requests. Inspect your file folder after each execution to verify that all the GIFs have been freshly downloaded. Turn in your three scripts and the three output sets, for example:

```
Elapsed time: 10.28958823
3110197 bytes downloaded
```

To get the number of bytes in a single flag, you can call **len(req.content)**, where **req** is the request object returned by **requests.get**.

FAQs

Q. Why can't I import requests?
A. Probably because you haven't installed it yet. It's a third-party module.

Q. Aren't we just doing the same thing over and over again?
A. Indeed, that's the point here. The thing that varies is whether we are using concurrency, and what type of concurrency we use. The goal is to have experience with concurrency.

Q. When I go to the CIA flag web site in a browser, I get an error.
A. You can only access each flag file directly with a complete URL that includes the name of each flag's GIF file.

Q. I missed class when you covered this. Can you help?
A. Not much. I will give you *one hint*: the code for the three versions is nearly identical. You should use a form of **map** in all 3 versions. Don't go off the deep end using other methods in the thread and process code other than the executor's **map** method. And for the process version, don't forget to put the driver code under a **if __name__ == '__main__':** guard.

Q. I don't like all of these flag files cluttering up my development folder.
A. Place them in a **flags** subfolder. It makes it easier to grade them as well.

Q. Don't you think Exercise 10.A gave us a huge head start on this one?
A. You're welcome.