

Project description:

=====

SGX publishes derivative data daily at the address below.

http://www.sgx.com/wps/portal/sgxweb/home/marketinfo/historical_data/derivatives/

Design a job to download the following files daily from the above website.

- 1) WEBPXTICK_DT-*.zip
- 2) TickData_structure.dat
- 3) TC_*.txt
- 4) TC_structure.dat

Requirements:

=====

- 1) It should be written in python and run like usual Linux commands, i.e. accepting command line options or even config file.
- 2) It should be able to download both historical files (files not on today) and today's file based on user's instructions.
- 3) Logging must be implemented.

Use logging module provided by python, which can provide flexible logging configurations, e.g. some messages are logged to both stdout and file and some to file only.

Make decisions on what messages/levels to log by yourself. The logs should help to debug/resolve issues.

- 4) The recovery plan should be considered. For example, you may ask yourself the following questions:

If the downloading failed on one day or on some days how do you redownload the missed file(s)?

Is the redownloading automatic or it requires manual intervention?

The website only lists the recent files. Is it possible to download older files?

Address any of your concerns in your design.