**ahrefs\_scraper**

3 scripts with steps:

* The first script will take input from input.xls and will generate the output file output-1.xls
* The second script will take input from input.xls and will generate the output file output-2.xls
* Then merge output-1.xls and output-2.xls to one file >> output-3.xls
* The third script will take output-3.xls and will generate the output file output-4.xls

**Video demo**

Youtube link: <https://youtu.be/wS5VYGrA1i4>

**Guides on how to run your script**

################################################

1. Install Python and firefox browser on your windows system

* Link download Python: <https://www.python.org/downloads/>, please choose new version 3.9.x
* Link download Firefox browser: <https://www.mozilla.org/en-US/firefox/new/>

################################################

1. Install package for your projects

* Open Command Line: Hold down Shift and right click on your project folder => Click to choose "Open PowerShell window here"
* Install package: Copy "pip install -r requirements.txt" and paste to your command line => Press Enter

################################################

1. Config settings file

* Open settings.py file in your folder project and choose edit with your editor.
* Fill your Ahrefs account
* Fill others (optional)

################################################

1. Run your script

* Open Command Line: Hold down Shift and right click on your project folder => Click to choose "Open PowerShell window here"
* Run script 1: Copy "python script\_1.py" and paste to your command line => Press Enter
* After complete script 1, run script 2: Copy "python script\_2.py" and paste to your command line => Press Enter
* After complete script 2, run script 3: Copy "python script\_3.py" and paste to your command line => Press Enter

PS: If you completed steps above, you don't need to repeat it.

**NOTES:**

* Run script 1 need to have file "input.xlsx"
* Run script 2 need to have file "input.xlsx" and "output-1.xlsx"
* Run script 3 need to have file "output-3.xlsx" and "blacklist keywords.txt"

**How to deploy script on Linux Server**

Request: Ubuntu 18.04.5 LTS

1. Install dependencies. sudo apt-get update sudo apt-get install -y unzip openjdk-8-jre-headless xvfb libxi6 libgconf-2-4 git-all firefox python3 python3-dev python3-pip python3-venv libxml2-dev libxslt1-dev zlib1g-dev libffi-dev libssl-dev
2. Download ahrefs\_scraper project: <https://github.com/luantechrius/ahrefs_scraper> and put in your server
3. Go to ahrefs\_scraper folder and install packages: pip3 install -r requirements.txt
4. Create a file run\_script.sh (nano run\_script.sh) and add this code: nohup python3 script\_1.py nohup python3 script\_2.py nohup python3 script\_3.py
5. Change permission file and run script: chmod +x run\_script.sh ./run\_script.sh