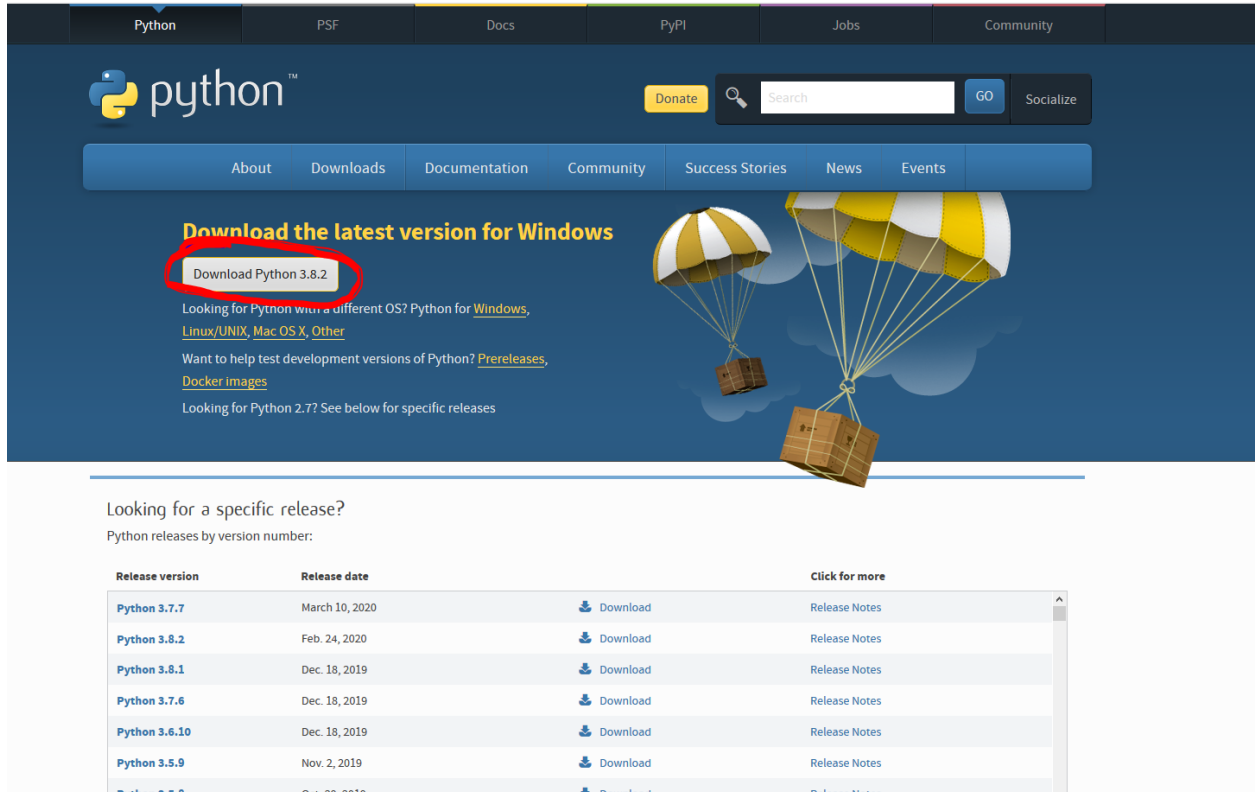


Skroutz Web Scrap Installation Guides for Windows OS

1) Install Python 3 using the following URL:

<https://www.python.org/downloads/>

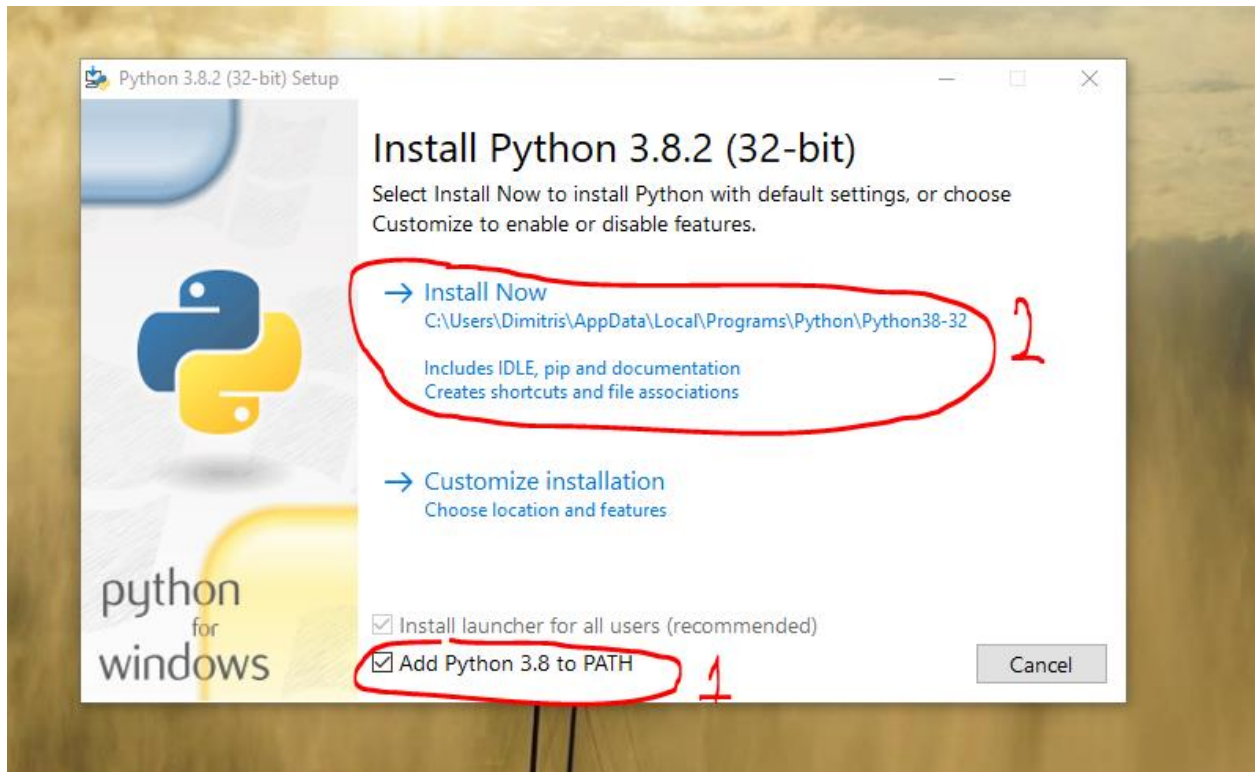


The screenshot shows the Python.org website. The main navigation bar includes links for Python, PSF, Docs, PyPI, Jobs, and Community. Below this, there's a search bar and a 'Donate' button. The main content area features a large banner with the text 'Download the latest version for Windows' and a button labeled 'Download Python 3.8.2', which is circled in red. To the right of the banner is an illustration of two parachutes. Below the banner, there's a section titled 'Looking for a specific release?' with the text 'Python releases by version number:'. This section contains a table with the following data:

Release version	Release date		Click for more
Python 3.7.7	March 10, 2020	Download	Release Notes
Python 3.8.2	Feb. 24, 2020	Download	Release Notes
Python 3.8.1	Dec. 18, 2019	Download	Release Notes
Python 3.7.6	Dec. 18, 2019	Download	Release Notes
Python 3.6.10	Dec. 18, 2019	Download	Release Notes
Python 3.5.9	Nov. 2, 2019	Download	Release Notes
Python 3.5.8	Oct. 29, 2019	Download	Release Notes

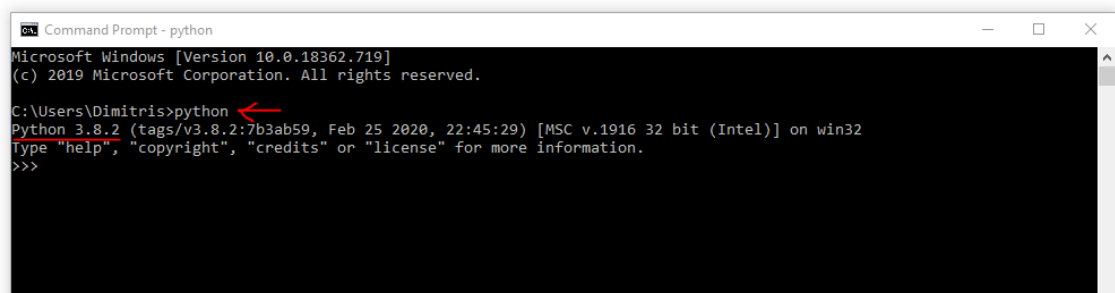
2) Find the downloaded .exe file and double click to run installation.

3) Click “Add Python 3.X to PATH and then select “Install Now” as shown below.



In case you want to check if Python is successfully installed, open a new Command Prompt terminal(Start -> type “cmd” ->Open Command Prompt) and type: “python”

If everything is correct, you will see the Python version you downloaded as shown below.

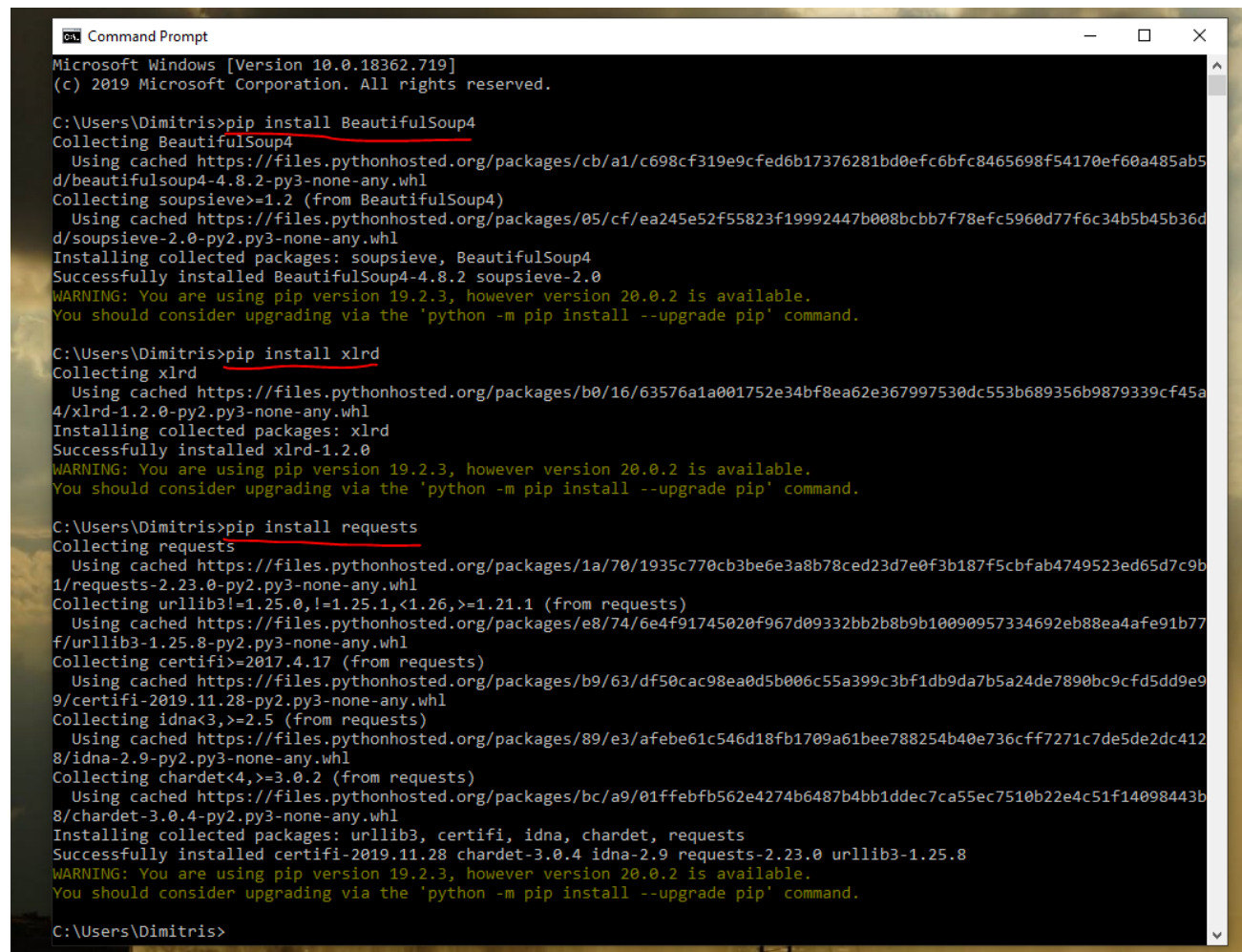


Exit Python terminal by pressing CTRL-Z.

4) After installation is completed, open a Command Prompt terminal(Start -> type “cmd” ->Open Command Prompt) to install the required Python libraries

5) Type the following commands:

```
$ pip install BeautifulSoup4
$ pip install xlrd
$ pip install requests
```

A screenshot of a Windows Command Prompt window. The title bar reads "Command Prompt". The window content shows the execution of three pip install commands. The first command is "pip install BeautifulSoup4", which shows the collection and installation of BeautifulSoup4-4.8.2 and soupsieve-2.0. The second command is "pip install xlrd", which shows the collection and installation of xlrd-1.2.0. The third command is "pip install requests", which shows the collection and installation of requests-2.23.0 and its dependencies: urllib3-1.25.8, certifi-2019.11.28, idna-2.9, and chardet-3.0.4. Each installation step includes a warning about using pip version 19.2.3 instead of 20.0.2 and a suggestion to upgrade using "python -m pip install --upgrade pip". The prompt path is "C:\Users\Dimitris>".

```
Microsoft Windows [Version 10.0.18362.719]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\Dimitris>pip install BeautifulSoup4
Collecting BeautifulSoup4
  Using cached https://files.pythonhosted.org/packages/cb/a1/c698cf319e9cfed6b17376281bd0efc6bfc8465698f54170ef60a485ab5d/beautifulsoup4-4.8.2-py3-none-any.whl
Collecting soupsieve>=1.2 (from BeautifulSoup4)
  Using cached https://files.pythonhosted.org/packages/05/cf/ea245e52f55823f19992447b008bcb7f78efc5960d77f6c34b5b45b36dd/soupsieve-2.0-py2.py3-none-any.whl
Installing collected packages: soupsieve, BeautifulSoup4
Successfully installed BeautifulSoup4-4.8.2 soupsieve-2.0
WARNING: You are using pip version 19.2.3, however version 20.0.2 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.

C:\Users\Dimitris>pip install xlrd
Collecting xlrd
  Using cached https://files.pythonhosted.org/packages/b0/16/63576a1a001752e34bf8ea62e367997530dc553b689356b9879339cf45a4/xlrd-1.2.0-py2.py3-none-any.whl
Installing collected packages: xlrd
Successfully installed xlrd-1.2.0
WARNING: You are using pip version 19.2.3, however version 20.0.2 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.

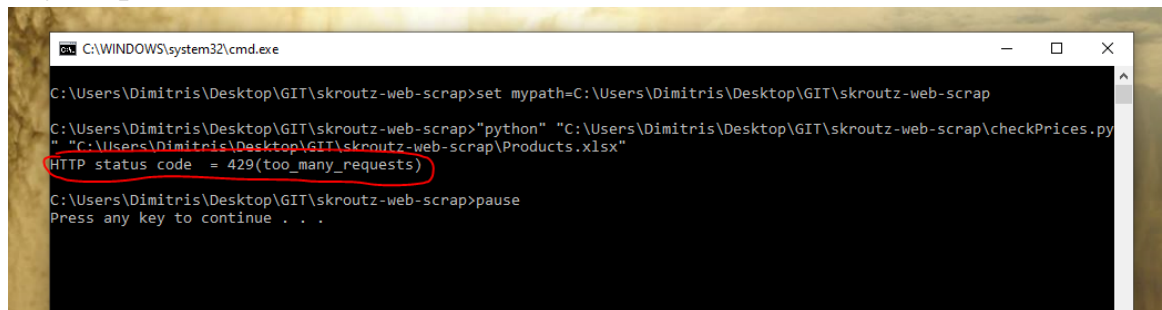
C:\Users\Dimitris>pip install requests
Collecting requests
  Using cached https://files.pythonhosted.org/packages/1a/70/1935c770cb3be6e3a8b78ced23d7e0f3b187f5cbfab4749523ed65d7c9b1/requests-2.23.0-py2.py3-none-any.whl
Collecting urllib3!>=1.25.0,!>=1.25.1,<1.26,>=1.21.1 (from requests)
  Using cached https://files.pythonhosted.org/packages/e8/74/6e4f91745020f967d09332bb2b8b9b10090957334692eb88ea4afe91b77f/urllib3-1.25.8-py2.py3-none-any.whl
Collecting certifi>=2017.4.17 (from requests)
  Using cached https://files.pythonhosted.org/packages/b9/63/df50cac98ea0d5b006c55a399c3bf1db9da7b5a24de7890bc9cfd5dd9e99/certifi-2019.11.28-py2.py3-none-any.whl
Collecting idna<3,>=2.5 (from requests)
  Using cached https://files.pythonhosted.org/packages/89/e3/afebe61c546d18fb1709a61bee788254b40e736cff7271c7de5de2dc4128/idna-2.9-py2.py3-none-any.whl
Collecting chardet<4,>=3.0.2 (from requests)
  Using cached https://files.pythonhosted.org/packages/bc/a9/01ffebfb562e4274b6487b4bb1ddec7ca55ec7510b22e4c51f14098443b8/chardet-3.0.4-py2.py3-none-any.whl
Installing collected packages: urllib3, certifi, idna, chardet, requests
Successfully installed certifi-2019.11.28 chardet-3.0.4 idna-2.9 requests-2.23.0 urllib3-1.25.8
WARNING: You are using pip version 19.2.3, however version 20.0.2 is available.
You should consider upgrading via the 'python -m pip install --upgrade pip' command.

C:\Users\Dimitris>
```

- 6) In order to run program, the files *checkPrices.py* *run_script.bat* and *Products.xlsx* must be placed in the same directory.
- 7) Then, double click *run_script.bat* in order to run the program.
- 8) When the program is successfully finished, the results with the required products are stored in a .csv file.

!!! IMPORTANT NOTICE !!!

When running the program you may see the following message without any output .csv file:



A screenshot of a Windows command prompt window. The title bar shows 'C:\WINDOWS\system32\cmd.exe'. The command prompt shows the following text:
C:\Users\Dimitris\Desktop\GIT\skroutz-web-scrap>set mypath=C:\Users\Dimitris\Desktop\GIT\skroutz-web-scrap
C:\Users\Dimitris\Desktop\GIT\skroutz-web-scrap>"python" "C:\Users\Dimitris\Desktop\GIT\skroutz-web-scrap\checkPrices.py"
"C:\Users\Dimitris\Desktop\GIT\skroutz-web-scrap\Products.xlsx"
HTTP status code = 429(too_many_requests)
C:\Users\Dimitris\Desktop\GIT\skroutz-web-scrap>pause
Press any key to continue . . .

Since Skroutz takes its protection measures against multiple HTTP requests, we can act as following:

Either run the program from a computer with a different IP or wait a few hours and rerun the program.

- 9) Open the .csv file preferably by Notepad++(or Notepad)
When opening the file with Excel sheets, greek shop names are not displayed correctly.