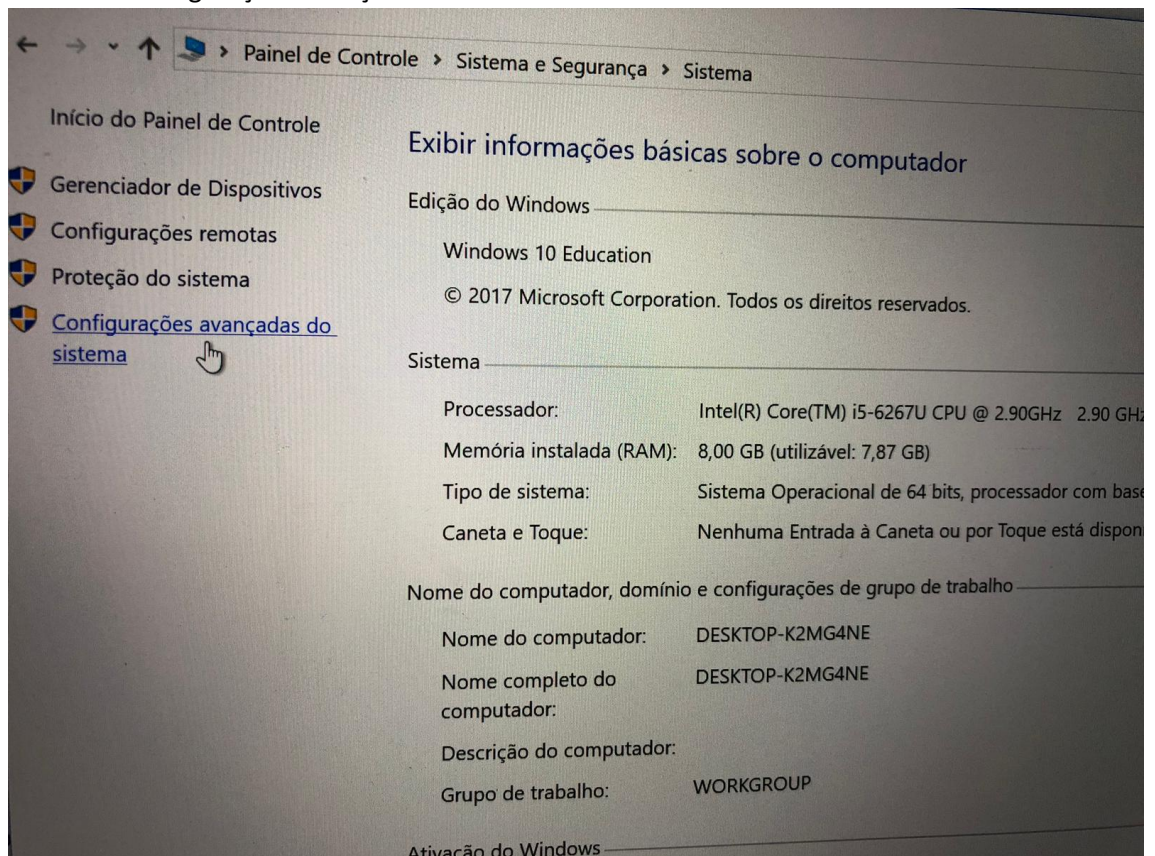


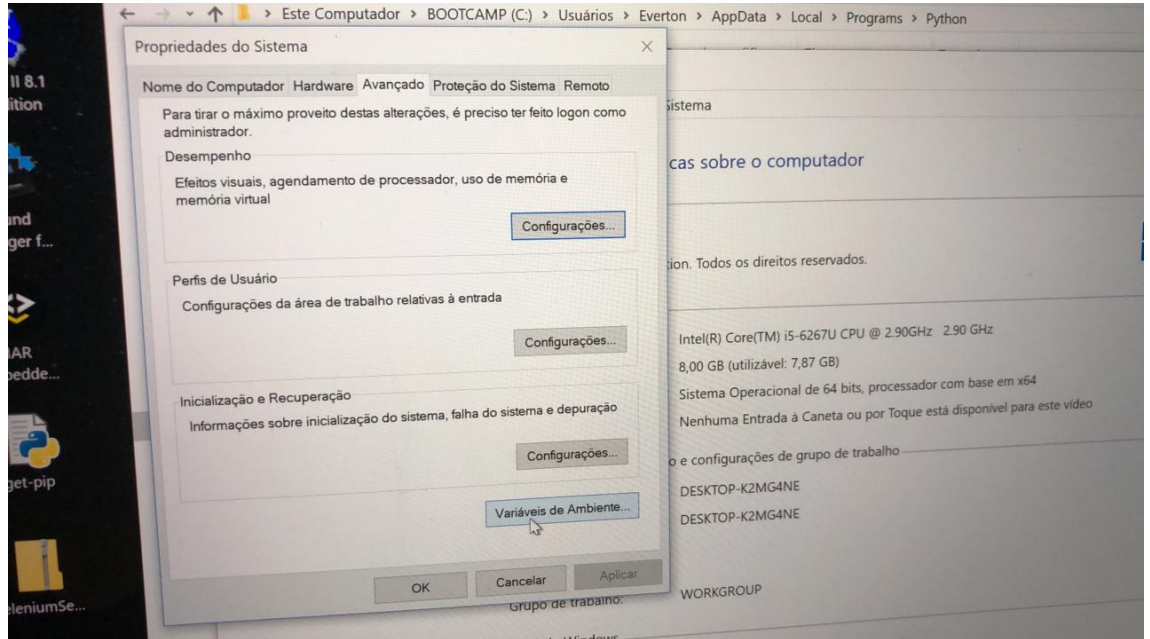
# Tutorial Instalação Python + PIP + Crawler

Windows:

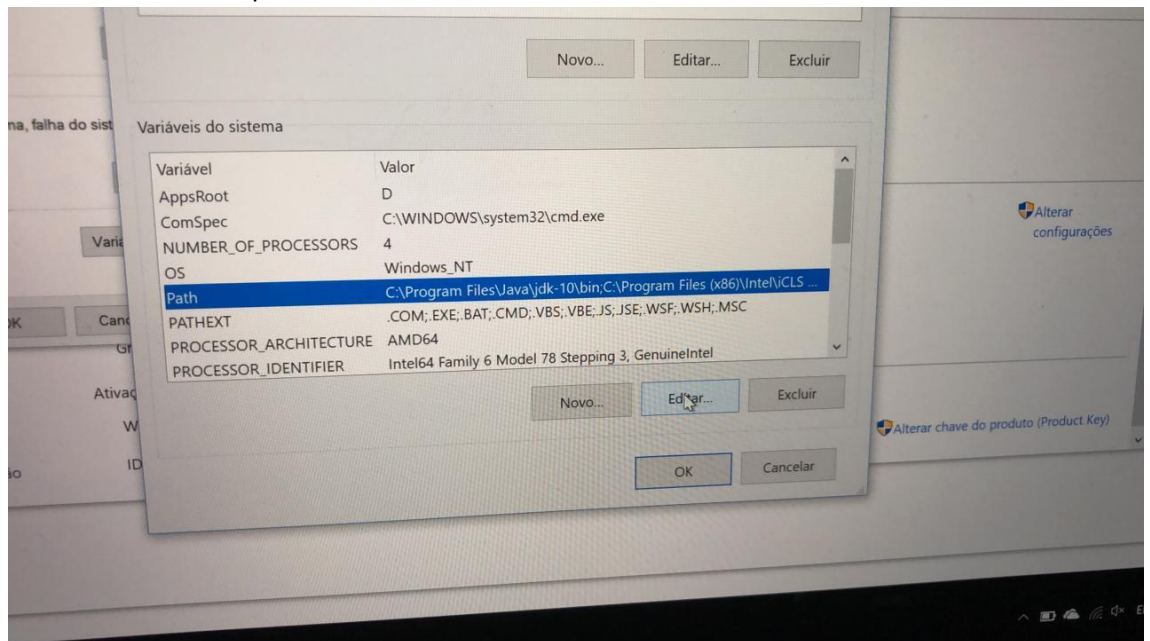
- 1- Baixar a última versão do Python em <https://www.python.org/ftp/python/3.7.2/python-3.7.2.exe>
- 2- Fazer a instalação e copiar o endereço em que o python foi instalado. Normalmente o endereço é C:\Users\USUARIO\AppData\Local\Programs\Python\Python37-32
- 3- Após a instalação, deve configurar o cmd para reconhecer o comando python. Para isso, entre no Painel de Controle, Sistema e Segurança e por fim, Sistema.
- 4- Clicar em Configurações avançadas do sistema:



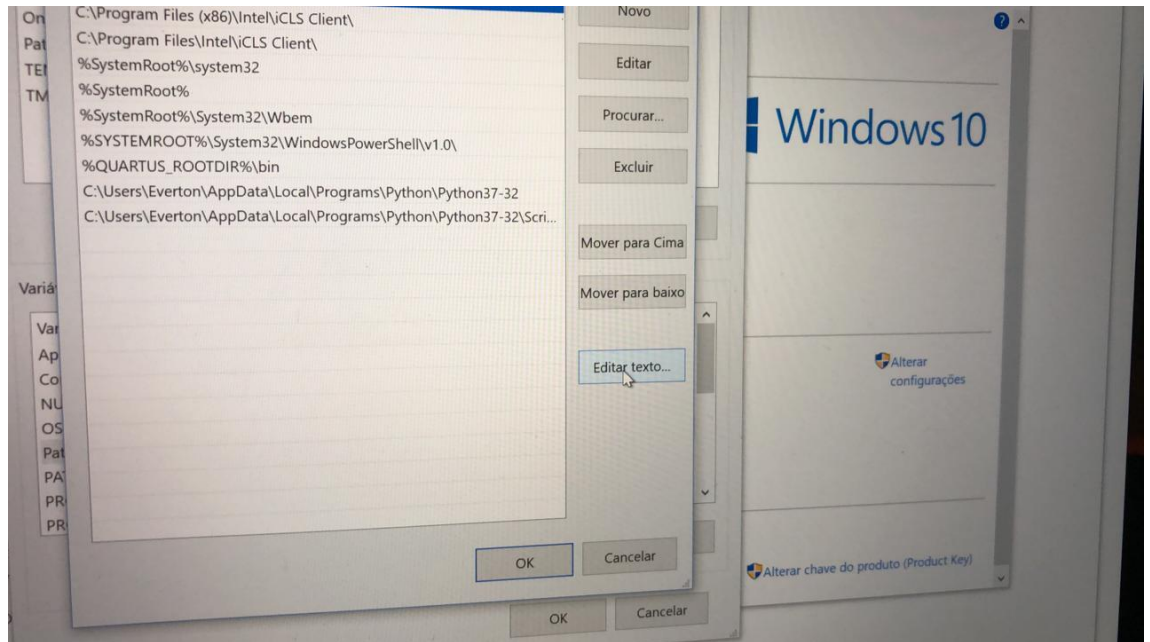
5- Clicar em Variáveis de Ambiente...:



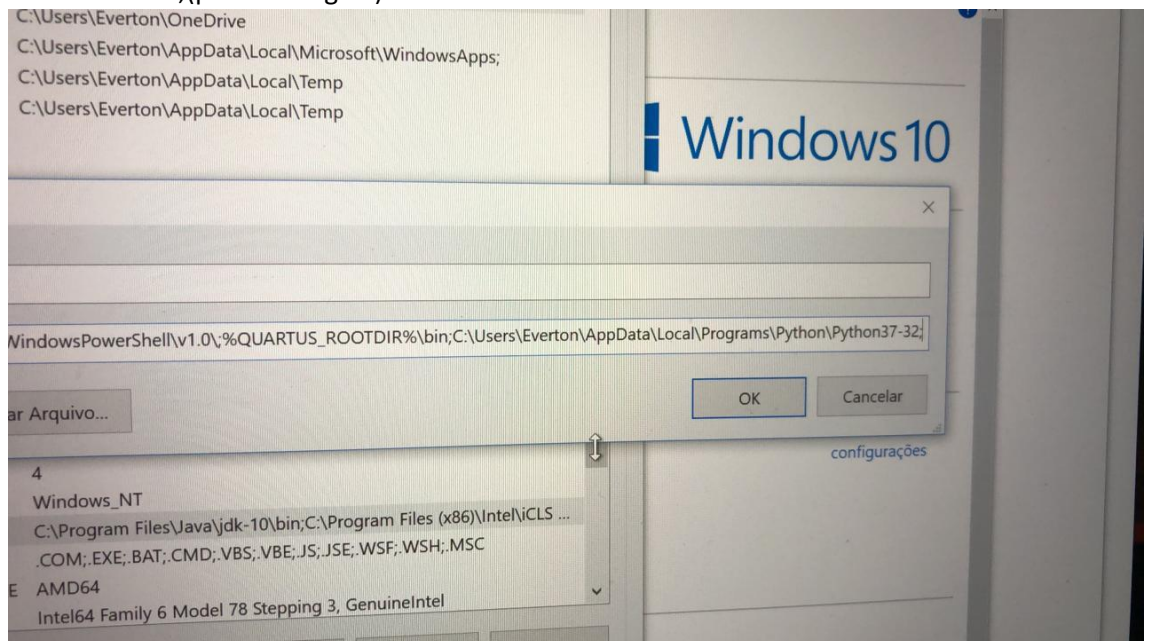
6- Selecionar Path e depois em Editar:



7- Clicar em Editar Texto:

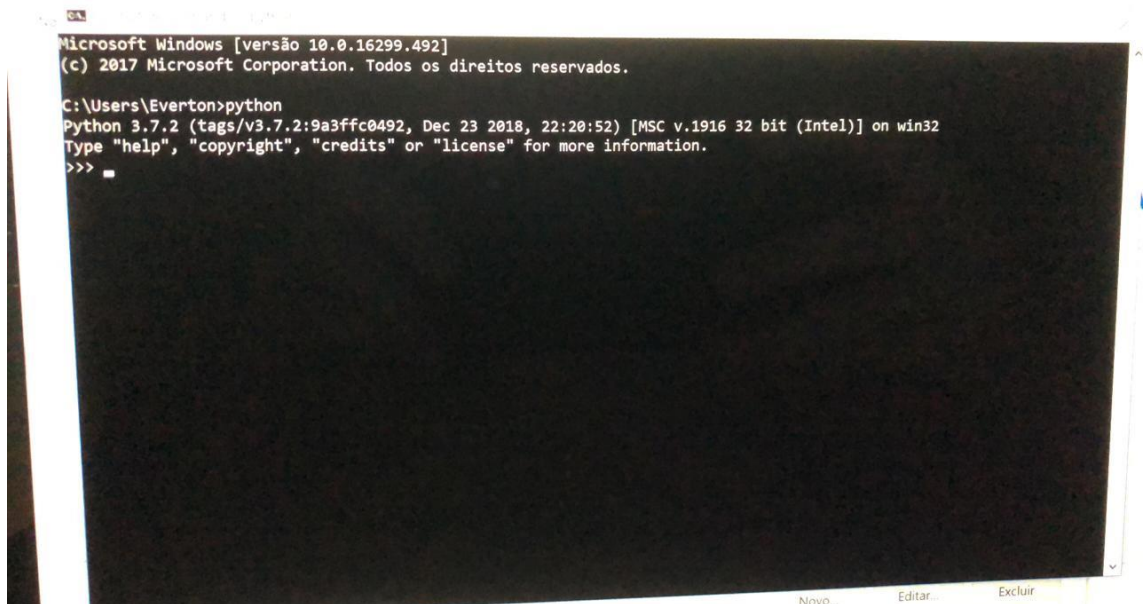


8- Ao final da linha, após o ; adicionar o endereço em que o Python foi instalado e terminar com ;(ponto e virgula) e confirmar com o OK:





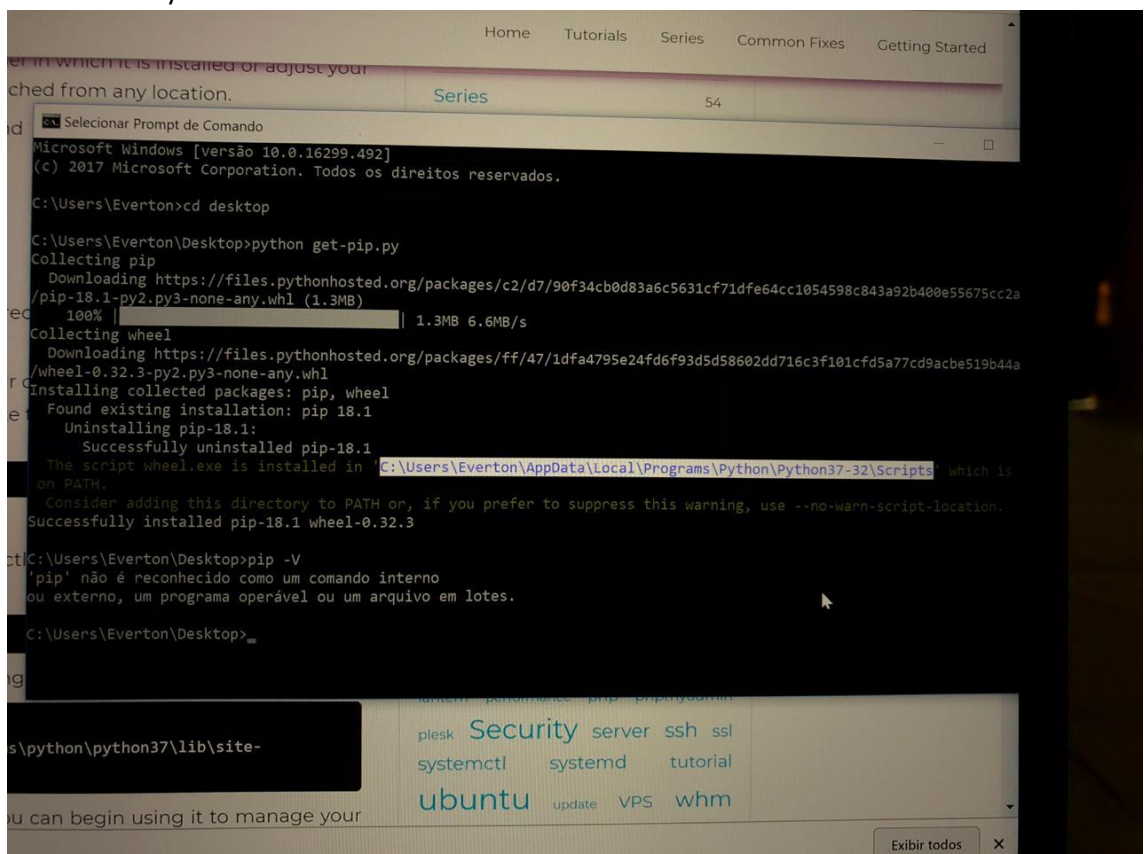
- 9- Abrir o cmd e digitar Python. Esta é a tela esperada se o processo anterior foi realizado corretamente:



```
Microsoft Windows [versão 10.0.16299.492]
(c) 2017 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Everton>python
Python 3.7.2 (tags/v3.7.2:9a3ffc0492, Dec 23 2018, 22:20:52) [MSC v.1916 32 bit (Intel)] on win32
Type "help", "copyright", "credits" or "license" for more information.
>>> _
```

- 10- Agora deve-se instalar o PIP (PIP é um gerenciador de pacotes por linha de comando para Python). Salve o arquivo get-pip.py contido neste .zip em uma pasta conhecida.
- 11- Abra o cmd e navegue para a pasta citada acima.
- 12- Digite python get-pip.py. O PIP será instalado e será informado no cmd a pasta em que o mesmo foi instalado. Deve-se copiar este endereço para fazer o mesmo processo feito com o Python:



```
Microsoft Windows [versão 10.0.16299.492]
(c) 2017 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Everton>cd desktop
C:\Users\Everton\Desktop>python get-pip.py
Collecting pip
  Downloading https://files.pythonhosted.org/packages/c2/d7/90f34cb0d83a6c5631cf71dfe64cc1054598c843a92b400e55675cc2a/pip-18.1-py2.py3-none-any.whl (1.3MB)
    100% |#####| 1.3MB 6.6MB/s
Collecting wheel
  Downloading https://files.pythonhosted.org/packages/ff/47/1dfa4795e24fd6f93d5d58602dd716c3f101cfd5a77cd9acbe519b44a/wheel-0.32.3-py2.py3-none-any.whl
Installing collected packages: pip, wheel
Found existing installation: pip 18.1
Uninstalling pip-18.1:
  Successfully uninstalled pip-18.1
The script wheel.exe is installed in 'C:\Users\Everton\AppData\Local\Programs\Python\Python37-32\Scripts' which is on PATH.
Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
Successfully installed pip-18.1 wheel-0.32.3

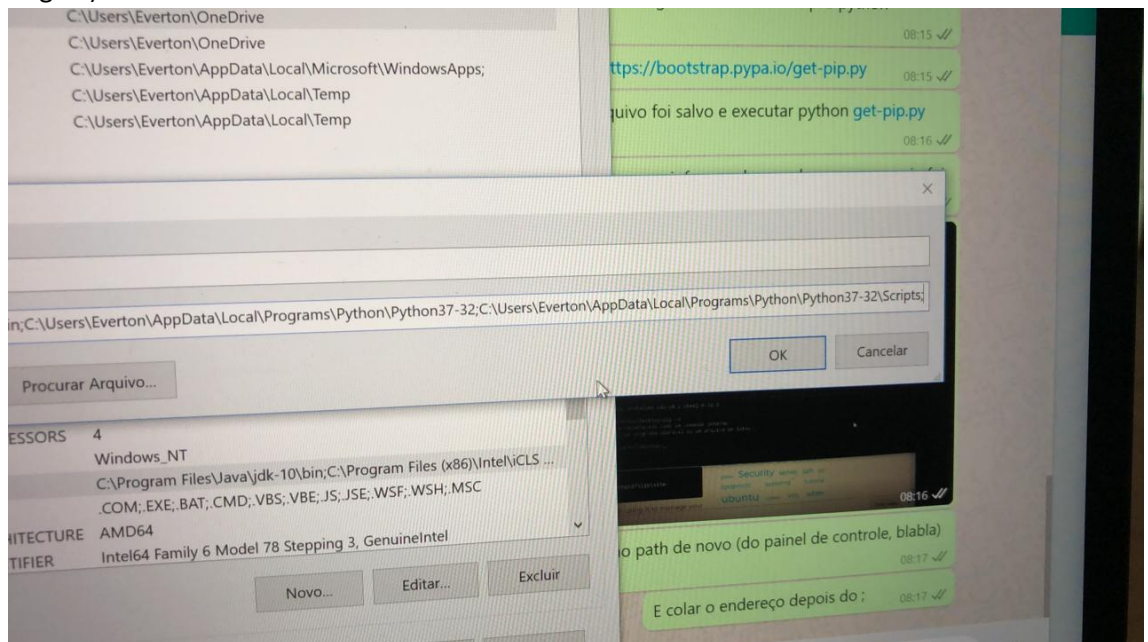
C:\Users\Everton\Desktop>pip -V
'pip' não é reconhecido como um comando interno
ou externo, um programa operável ou um arquivo em lotes.

C:\Users\Everton\Desktop>_
```

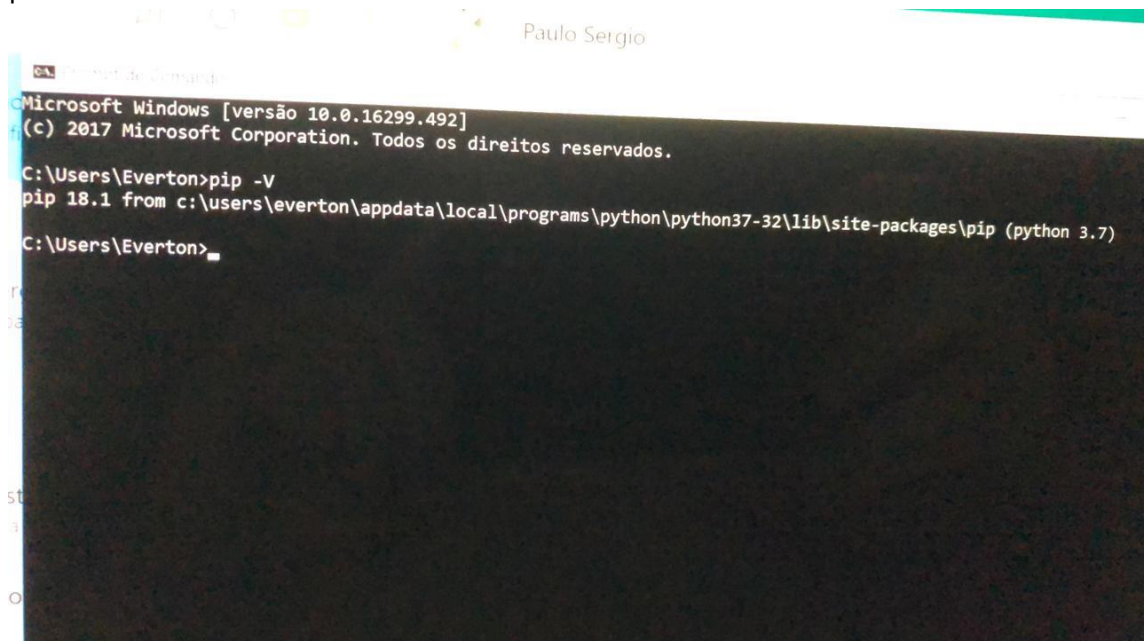
Search results for 'python\python37\lib\site-':

- Security server ssh ssl
- systemctl systemd tutorial
- ubuntu update vps whm

- 13- Seguir os passos 4 até 7 e adicionar ao final do ultimo ; (ponto e virgula), o endereço em que o PIP foi instalado e informado no cmd e terminar novamente com o ;(ponto e virgula):

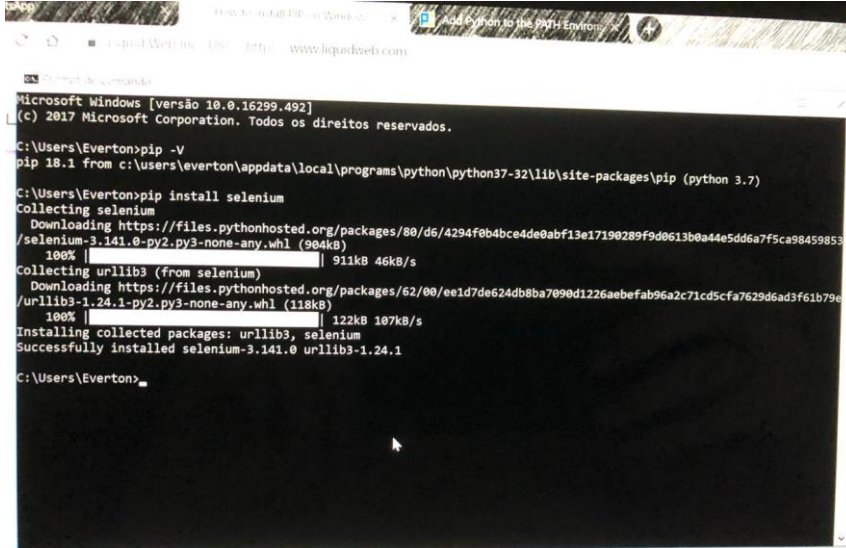


- 14- Após a confirmação, abrir um novo cmd e digitar pip -V. Esta é a tela esperada caso o processo tenha sido feito corretamente:



- 15- Agora deve-se instalar os pacotes que foram usados no crawler.

16- Digitar no cmd pip install selenium:



```
Microsoft Windows [versão 10.0.16299.492]
(c) 2017 Microsoft Corporation. Todos os direitos reservados.

C:\Users\Everton>pip -V
pip 18.1 from c:\users\everton\appdata\local\programs\python\python37-32\lib\site-packages\pip (python 3.7)

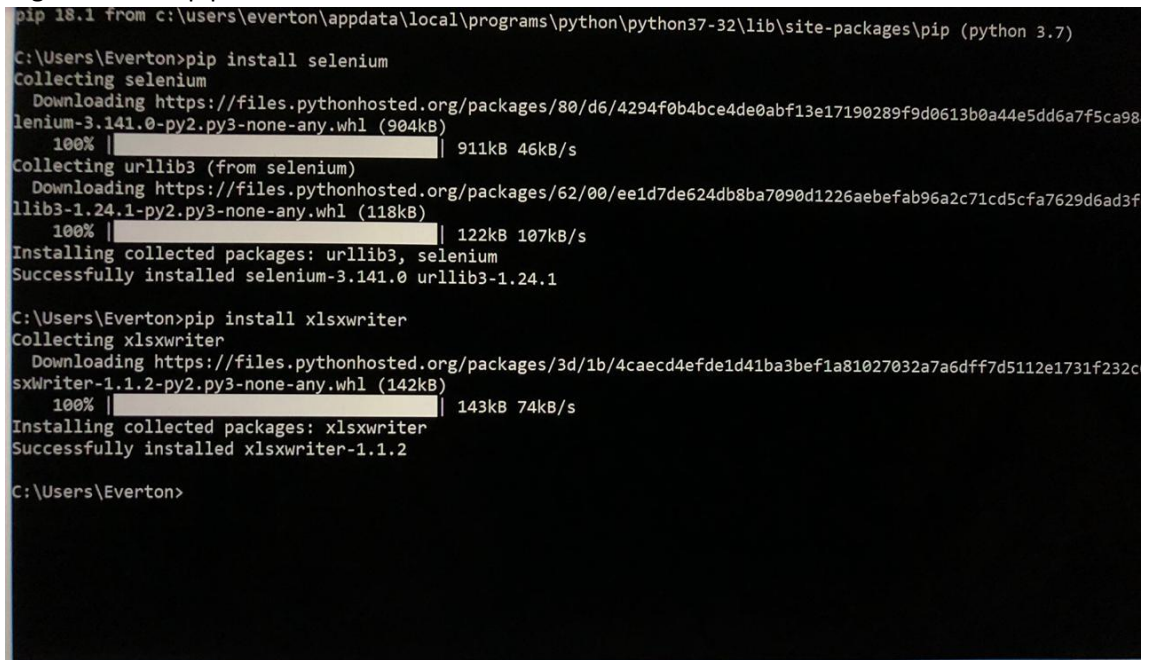
C:\Users\Everton>pip install selenium
Collecting selenium
  Downloading https://files.pythonhosted.org/packages/80/d6/4294f0b4bce4de0abf13e17190289f9d0613b0a44e5dd6a7f5ca98459853/selenium-3.141.0-py2.py3-none-any.whl (904kB)
    100% |#####| 911kB 46kB/s
Collecting urllib3 (from selenium)
  Downloading https://files.pythonhosted.org/packages/62/00/ee1d7de624db8ba7090d1226aebefab96a2c71cd5cfa7629d6ad3f61b79e/urllib3-1.24.1-py2.py3-none-any.whl (118kB)
    100% |#####| 122kB 107kB/s
Installing collected packages: urllib3, selenium
Successfully installed selenium-3.141.0 urllib3-1.24.1

C:\Users\Everton>
```

You can verify that Pip was installed correctly by opening a command prompt and entering the following command:

```
pip -V
```

17- Digitar no cmd pip install xlswriter:



```
pip 18.1 from c:\users\everton\appdata\local\programs\python\python37-32\lib\site-packages\pip (python 3.7)

C:\Users\Everton>pip install selenium
Collecting selenium
  Downloading https://files.pythonhosted.org/packages/80/d6/4294f0b4bce4de0abf13e17190289f9d0613b0a44e5dd6a7f5ca98459853/selenium-3.141.0-py2.py3-none-any.whl (904kB)
    100% |#####| 911kB 46kB/s
Collecting urllib3 (from selenium)
  Downloading https://files.pythonhosted.org/packages/62/00/ee1d7de624db8ba7090d1226aebefab96a2c71cd5cfa7629d6ad3f61b79e/urllib3-1.24.1-py2.py3-none-any.whl (118kB)
    100% |#####| 122kB 107kB/s
Installing collected packages: urllib3, selenium
Successfully installed selenium-3.141.0 urllib3-1.24.1

C:\Users\Everton>pip install xlswriter
Collecting xlswriter
  Downloading https://files.pythonhosted.org/packages/3d/1b/4caecd4efde1d41ba3bef1a81027032a7a6dff7d5112e1731f232c/xlswriter-1.1.2-py2.py3-none-any.whl (142kB)
    100% |#####| 143kB 74kB/s
Installing collected packages: xlswriter
Successfully installed xlswriter-1.1.2

C:\Users\Everton>
```

You should see output similar to the following:

```
pip 18.0 from
c:\users\administrator\appdata\local\programs\python\python37\lib\site-
packages\pip (python 3.7)
```

18- Após isso, todos os programas e pacotes necessários já estão instalados, bastando apenas rodar o crawler.

19- Navegar até o diretório do GitHub

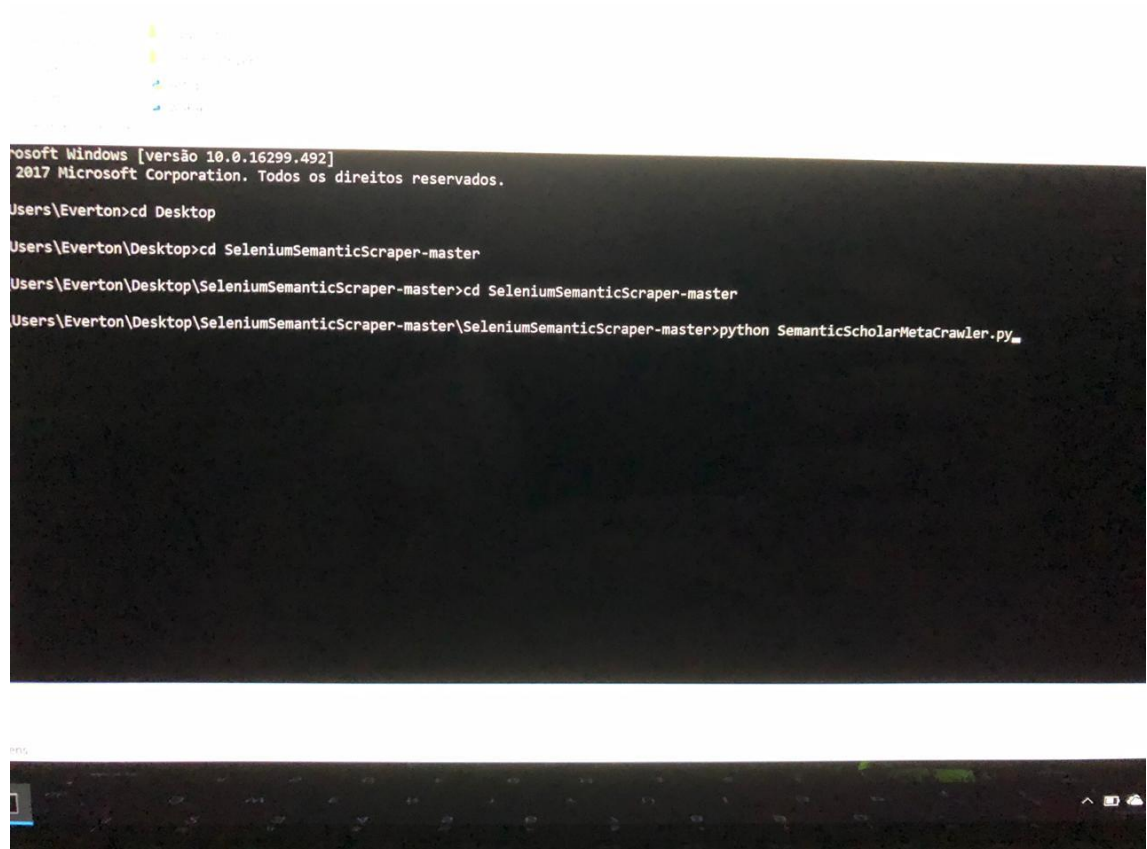
<https://github.com/EvertonCa/SeleniumSemanticScraper> e clicar em Clone or Download e clicar em Download ZIP.

20- Descompactar a pasta em um local conhecido.

21- Abrir um terminal cmd e navegar até a pasta descompactada



- 22- Executar o arquivo SemanticScholarMetaCrawler.py, digitando no cmd “python SemanticScholarMetaCrawler.py”:



```
Microsoft Windows [versão 10.0.16299.492]  
2017 Microsoft Corporation. Todos os direitos reservados.  
  
Users\Everton>cd Desktop  
  
Users\Everton\Desktop>cd SeleniumSemanticScraper-master  
  
Users\Everton\Desktop\SeleniumSemanticScraper-master>cd SeleniumSemanticScraper-master  
  
Users\Everton\Desktop\SeleniumSemanticScraper-master\SeleniumSemanticScraper-master>python SemanticScholarMetaCrawler.py
```

- 23- O programa solicitará sua pesquisa (as palavras chave de sua pesquisa) e a quantidade de páginas que gostaria de pesquisar (cada página corresponde entre 10 e 30 artigos).
- 24- A primeira execução deve pedir permissão de firewall. Aceitar e em caso de erro no programa, executar novamente.
- 25- Após o término do programa (uma página do Google Chrome abrirá e começará a ser atualizada automaticamente, até o fim da execução) uma pasta Files será criada e dentro desta pasta, outra com o termo pesquisado será criada. O arquivo excel de sua pesquisa estará nesta página.