



Instalación snsrape

versión útil



Instalar una versión de Python superior a 3.8

Recordando que vamos a usar ambientes virtuales controlados y necesitamos una versión de Python mayor a 3.8 (3.8.1 es mayor).

Sidenote: si ya tienes una versión superior, ir a diapositiva siguiente.

```
sudo apt update
sudo apt install software-properties-common
sudo add-apt-repository ppa:deadsnakes/ppa
sudo apt install python3.8
sudo apt -y install python3-pip
sudo apt remove --purge python3-apt
sudo apt autoclean
sudo apt install python3-apt
sudo apt install python3.9-distutils
```



Instalación Pipenv y necesarios.

```
$> pip3 install --user pipenv
```

```
$> mkdir tweets
```

```
$> cd tweets
```

```
$> PYTHON_BIN="$(python3.8 -m site --user-base)/bin" (notar aquí que es el python instalado, sustituir si es python3.9 o python3.10....)
```

```
$> PATH="$PATH:$PYTHON_BIN"
```

```
$> source ~/.bashrc
```



Modificación de clase Tweets.py

`$>pipenv shell` *(aquí se crea, el ambiente virtual y se activará, verás entonces el nombre del ambiente en el prompt, en este caso tomará el nombre de la carpeta donde estés)*

`$(tweets)>pipenv install git+https://github.com/JustAnotherArchivist/snsrape@master#egg=snsrape`

Ahora, buscaremos el archivo tweets.py de la librería y lo reemplazaremos por el archivo tweets.py que está en Moodle.



Modificación de clase Tweets.py

```
(tweets)$> ls ~/.local/share/virtualenvs
```

El ambiente tiene una clave aleatoria, saldrá en la lista algo parecido a “tweets-heqVa2Z”, listaremos el contenido

```
(tweets)$> ls  
~/.local/share/virtualenvs/tweets-heqVa2Z/lib/python3.8/site-packages/snsrape/modules/
```

Modificar la ruta si la versión de python no es 3.8



Modificación de clase Tweets.py

Identificar dónde está el archivo descargado de moodle/git y cargado en la instancia o donde se vaya a correr el proceso, suponiendo que está en home (~), haremos la siguiente línea:

```
(tweets)$> cp ~/tweets.py  
~/.local/share/virtualenvs/tweets-heqVa2Z/lib/python3.8/site-packages/snsrape/modules/
```

Esto reemplazará el archivo existente y ¡Estamos listos para “scrapear” twitter!



Ejemplos

Descarga por hashtag

```
snsrape --jsonl --max-results 100 twitter-hashtag tbt > busqueda_hashtag.txt
```

Descarga por usuario

```
snsrape --jsonl --max-results 100 twitter-user elonmusk > busqueda_usuario.txt
```

Descarga por query

```
snsrape --jsonl --max-results 100 twitter-search 'Ukraine OR ucrania since:2022-03-01 until:2022-03-28' > busqueda_query.txt
```



Referencias.

<https://github.com/JustAnotherArchivist/snsrape>

<https://github.com/JustAnotherArchivist/snsrape/issues/407#issuecomment-1073035041>