

Overview: The Twitter advanced search page does not come with all the results already embedded in it, but instead downloads 15-30 tweets on demand each time the user scrolls to the bottom of the page. These come in the form of a JSON file, which is requested with a standard URL. Once one has one request URL from a search, one can predict the next URL by replacing initial URL's first tweet ID with the last tweet ID returned by that URL. One then uses the updated request URL to get another JSON file from Twitter, find the last tweet ID, amend the request URL again, etc. This URL also contains the final tweet ID to be returned by the entire search, which cannot be predicted algorithmically, so one must get at least one request URL manually.

Steps to use the software:

1. Install Python 3. It must be a 3.x version of Python. The version I've been using on both Mac and Windows is version 3.4.1.
2. Open Terminal and type "python" and press Enter. If it says "command not found" you need to add the Python directory to your PATH. This tutorial might help or you can just google it:  
[http://www.tutorialspoint.com/python/python\\_environment.htm](http://www.tutorialspoint.com/python/python_environment.htm)
3. Once Python is installed and available from any directory, download the necessary files and put them together in some folder. These are the 3 .py files on Github (these are probably synced to your computer already).
4. I wouldn't worry about getting the actual tweet contents yet, since that's trivial (and fast) once we have the IDs stored. Getting the IDs is going to be the most time-consuming part. So, to store the IDs, this is what you need to do
  - a. Go to <https://twitter.com/search-advanced> in Chrome and enter the word, set the language to English, and put in the dates. I've been going from Jan 1, 2011 to Sept 30, 2015, but depending on how many results are returned (it varies from hundreds to hundreds of thousands) we may want to break it into individual years or even smaller increments instead of doing one giant search.
  - b. In the Chrome menubar, go to View → Developer → Developer Tools. A window will appear to the side. If it already seems to show many results, click the button to clear the output (the circle with the slash through it next to the red circle) and refresh the page.
  - c. Scroll down the Twitter page until you reach the bottom and it has to load more.
  - d. In the Developer Console window, there is a search box. Type "timeline?vert" into the search box.
  - e. If you don't see any results that look like timeline?vertical..., then try step C again.
  - f. Otherwise, click the first timeline?vertical... result. A bunch of information will appear to the right. Look for the **Request URL** and copy it to the clipboard (Yes, this means that those first ~30 tweets are excluded from the search. I don't think this is a problem).
  - g. Open Terminal, and navigate to the directory where you saved the python files
    - i. In Terminal, "cd X" changes your directory to X if X is a full path or a subdirectory of the current directory. Typing part of a name and then pressing tab completes the name to save typing. Again, there are many helpful tutorials online about using the Mac OS X Terminal.
  - h. Enter the command "python *"YOUR-MAILED-REQUEST-URL" OUTPUT-FILE-NAME"*
    - i. Note that the URL must have quotes around it, while the entire command must NOT have quotes around it. I show it in quotes here only for the sake of legibility.
    - ii. Output filename can be whatever.
  - i. Press Enter and it should run.
  - j. It will run until it has returned all the results. It will say "DONE" when it's done. Until then, it will update the number of results returned so far every couple minutes.

[illegible]