iConference 2013 Big Social Data Workshop Hands-On Activity

**What you need** (*these steps are for the Mac platform*)

* + Twitter account credentials – a user name and password combination
  + A browser plugin to format JSON data for easy reading
    - For Firefox – install JSONView 0.7 by going to Tools > Add-ons > Extensions and searching for “jsonview”
    - For Safari – install JSON Formatter 1.1 from <https://github.com/rfletcher/safari-json-formatter>

**Exercise 1 – Pulling twitter data using a web browser**

In this exercise, we will use an ordinary web browser to make data requests to the twitter API. The results will be shown as a human-friendly presentation of the JSON encoded data returned from twitter.

1. Use your browser to navigate to <https://dev.twitter.com/docs/api/1/get/search>
2. Find the “Parameters” section of the page and examine the various options
3. Find the “Example Request” section of the page and copy the URL next to the word GET
4. Paste the URL into the address field of your browser and go to the page
5. You should see results similar to the twitter example (Safari users may have to press the “Toggle Formatting” button in the upper right of the page
6. *Leave this browser tab open for use in the next exercise*

**Exercise 2 – Experimenting with search query construction**

This exercise will introduce twitter’s advanced query parameters and show how to use the twitter web site to help format your data requests in subsequent exercises.

1. Navigate to twitter’s search interface at <http://www.twitter.com/search>
2. Click the “operators” link just below the query entry field next to “Tip:” and make note of the options available then close the dialog box.
3. Click the “advanced search” link to navigate to the advanced search form.
4. Enter a query of interest using some the available fields then click the “Search” button.
5. When you arrive at the results page, copy everything *to the right of the question mark (starting with “q=”)* that appears in your browser’s address field.
6. Paste the portion of the URL you just copied into the browser tab that you left open from exercise 1 replacing the existing q=… string and go to the page. You should see search results.
7. Repeat steps 4 through 6 to experiment with various search capabilities.

**Exercise 3 – Transitioning to the command line**

We are now ready to capture twitter data to a disk file for safekeeping and future analysis. The exercise relies upon the cURL command line utility this is part of Linux and Mac OS systems.

1. Open the Terminal application – Applications > Utilities > Terminal
2. Try using the cURL utility by typing curl google.com and pressing return. You should see something like:  
   <HTML><HEAD><meta http-equiv="content-type" content="text/html;charset=utf-8">  
   <TITLE>301 Moved</TITLE></HEAD><BODY>  
   <H1>301 Moved</H1>  
   The document has moved  
   <A HREF="http://www.google.com/">here</A>.  
   </BODY></HTML>
3. Return to <https://dev.twitter.com/docs/api/1/get/search> and copy the example request again as you did in exercise 1.
4. Returning to your terminal window, type curl followed by the URL now on your clipboard by pasting. The result should look like this:  
   curl http://search.twitter.com/search.json?q=blue%20angels…  
   Then press return. You should see your results as unformatted JSON data.
5. Look for the search term “blue angels” by using the Terminal program’s find feature Edit > Find > Find…

**Exercise 4 – Capturing filtered streaming data**

In this exercise we will be using a streaming API to collect tweets related to a set of filter terms. We will first view the results on the screen then learn to capture the data to a disk file. (see: https://dev.twitter.com/docs/api/1.1/post/statuses/filter)

1. Open the Terminal application – Applications > Utilities > Terminal
2. Enter curl --user *user name*:*password* -X POST -d 'track=milk,honey' <https://stream.twitter.com/1.1/statuses/filter.json>  
   and press enter.
3. Stop the data transfer by pressing control-c on the keyboard.
4. Look for the tracking terms you used in step 2 by using the Terminal program’s find feature Edit > Find > Find…
5. Now capture data to disk by entering curl -- user *user name*:*password* -X POST -d 'track=milk,honey' <https://stream.twitter.com/1.1/statuses/filter.json> > mydata.txt  
   and press enter.
6. After observing the transfer statistics for a while, stop the process by pressing control-c.
7. Type pwd and press enter to see the path where your data has been written.
8. Open the results file in any text editor to verify the capture process.