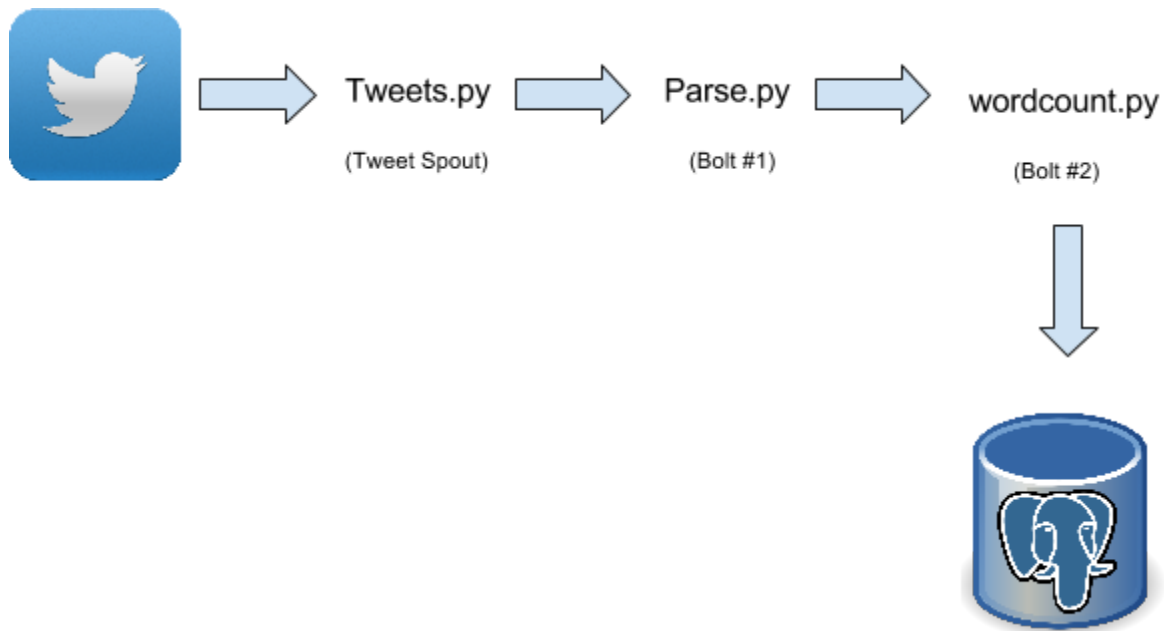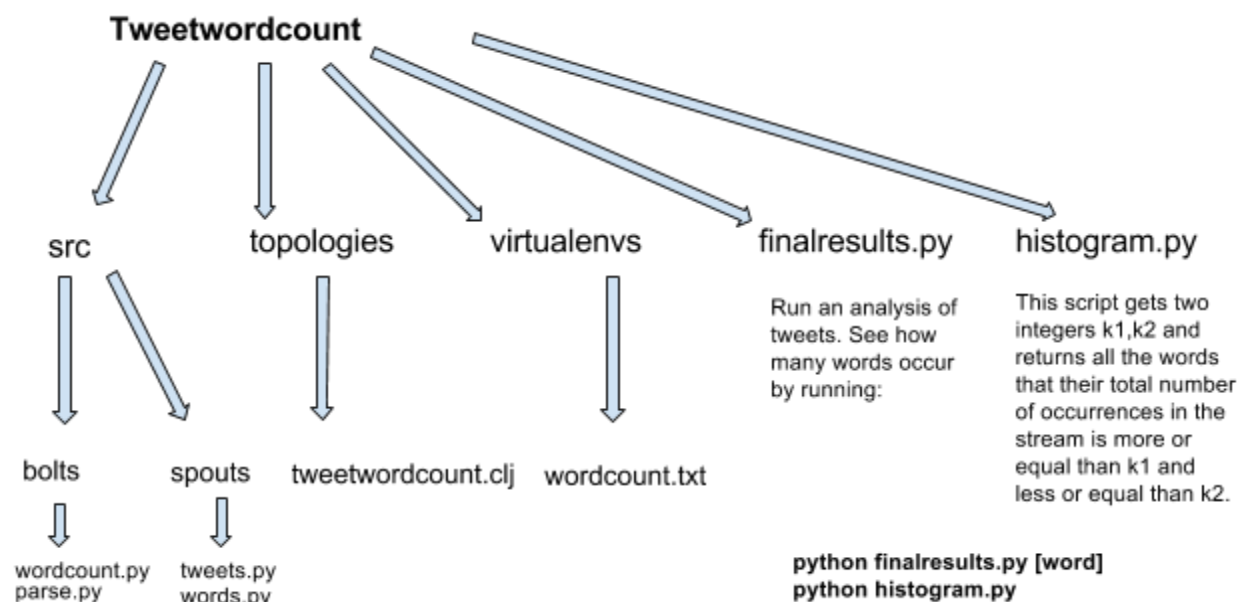# Twitter Application Documentation

This application accesses the twitter stream from twitter with the user's API key and other credentials, takes the twitter data, which is essentially sentences or lines of thought and computes the number of words (or a specific word) that occurred in the tweets processed.

Directory & File Structure:

src → **bolts** → wordcount.py parse.py

src → **spouts** → tweets.py words.py

topologies → tweetwordcount.clj

virtualenvs → wordcount.txt

finalresults.py — Run an analysis of tweets. See how many words occur by running:

histogram.py — This script gets two integers k1,k2 and returns all the words that their total number of occurrences in the stream is more or equal than k1 and less or equal than k2.

```
python finalresults.py [word]
python histogram.py
```

How to Run the Application:

In order to run the application, clone this file on your EC2 instance. On your instance, make sure to have postgres, python 2.7.3, the last version of streamparse (as of 8/7/16), tweepy, virtualenv, psycopg2, and redis installed. Create a new project on streamparse called Tweetwordcount. This file name should not have a number on it as it might create an error. Move the files above to your streamparse project, renaming it to Tweetwordcount. To run finalresults.py and get a wordcount, type *python finalresults.py* followed by the word you are looking for. You can do the same for histogram.py. If you want to make the project your own, create an API with twitter and then fill in your access codes accordingly.