

# Exercise 2

W205, Spring 2018

Author: Jason Salter

This document will outline the setup, architecture, and running of an application designed to parse tweets into a postgres SQL database. The primary directory is 'exercise2', which contains:



## SETUP:

Start Postgres database with `/data/start_postgres.sh`  
`cd exercise2/exttweetwordcount/src/spouts`  
`vi tweets.py` and add twitter credentials from you account.

## INTERACTING WITH THE STORED TWEETS

- 1) `cd exercise2/exttweetwordcount`
- 2) sparse run
- 3) start a second cmd window, and `cd exercise2`
  - a) `python finalresults.py` (ENTER)
  - b) if search word is desired, enter it here. otherwise, (ENTER) again
  - c) \*\*\* please note, searching on a single quote can be escaped with a second single quote: `you're` becomes `you"re`
  - d) all inputs are converted to lower case, as are the tweets
- 4) `python histogram.py` (ENTER)
  - a) Enter minimum count (ENTER)
  - b) Enter maximum count (ENTER)

## TEARDOWN:

- 1) Go to the cmd window running sparse run and CTRL-C to quit.
- 2) in cmd: `psql -U postgres`
  - a) drop database `tweetcount`;
  - b) ensures a fresh run, prevents OID errors on re-run after AMI

restart.

c) \q to quit

3) logout (to root)

4)/data/stop\_postgres.sh