W205-1  Spring 2016

Alfred W. Arsenault

Instructions for Running Exercise 2:

1.  Clone the repository:

    git clone https://github.com/al-arsenault/exercise_2

2.  Ensure that Hadoop and postgresql are running on the target machine; e.g. by running the
    /root/start-hadoop.sh and /data/start_postgres.sh shell scripts.
3.  Ensure that all necessary support software is running; e.g., psycopg2, pip, etc.
4.  Create a Streamparse project called EX2Tweetwordcount using the command

    $sparse quickstart EX2Tweetwordcount

5.  Run the program by issuing the commands:
        $cd EX2Tweetwordcount
        $sparse run
6.  When the program has ended, the files finalresults.py and histogram.py can be used to analyze
    the results.