**Take Home:**

A directory on the local machine holds about 1000 files that are HTTP access logs in common log format. Each log file is from yesterday, contains 24 hours of data and is about 500MB. Write a script that will produce a minute­-by-­minute CSV that contains an aggregate count of the accesses organized by HTTP status code (200 and 404 in the example below). Below is the sample of the log data and expected output:

**Log Sample:**

123.65.150.10 - - [23/Aug/2010:03:50:58 +0000] "POST /wordpress3/wp-admin/admin-ajax.php HTTP/1.1" 200 2 "<http://www.example.com/wordpress3/wp-admin/post-new.php>" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10\_6\_4; en-US) AppleWebKit/534.3 (KHTML, like Gecko) Chrome/6.0.472.25 Safari/534.3"

143.33.150.10 - - [23/Aug/2010:03:50:59 +0000] "POST /wordpress3/wp-admin/admin-ajax.php HTTP/1.1" 200 2 "<http://www.example.com/wordpress3/wp-admin/post-new.php>" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10\_6\_4; en-US) AppleWebKit/534.3 (KHTML, like Gecko) Chrome/6.0.472.25 Safari/534.3"

99.33.110.10 - - [23/Aug/2010:03:51:02 +0000] "GET /wordpress3/wp-admin/not-exist.php HTTP/1.1" 404 2 "<http://www.example.com/wordpress3/wp-admin/post-new.php>" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10\_6\_4; en-US) AppleWebKit/534.3 (KHTML, like Gecko) Chrome/6.0.472.25 Safari/534.3"

99.33.110.10 - - [23/Aug/2010:03:51:05 +0000] "GET /wordpress3/wp-admin/not-exist.php HTTP/1.1" 200 2 "<http://www.example.com/wordpress3/wp-admin/post-new.php>" "Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10\_6\_4; en-US) AppleWebKit/534.3 (KHTML, like Gecko) Chrome/6.0.472.25 Safari/534.3"

**Expected Result:**

# 200

23/Aug/2010:03:50, 2

23/Aug/2010:03:51, 1

# 404

23/Aug/2010:03:51, 1