Raw Logs Processing

Nitish Rangarajan

November 22, 2017

Preprocess the logs into a csv dataset

The raw dataset was collected as a log from the servers. The log has two types of arrays namely Persistent_Storage and RoundInfo. [3] The avatar's history is stored in the Persistent_Storage. Each element in this array contains the information about the avatar during the observed timeframe. They contain 11 fields which are separated by commas. They are dummy, query_time, query_sequence_number, avatar_ID, guild, level, race, character_class, zone, dummy and dummy.

Create a regular expression that would process the logs.

```
import csv
import re
import os           # This is for os.listdir
import os.path      # This is for the other dir stuff.
import string       # Maybe for directory name cycling etc.
import time         # For timing how long it takes.

# VARIABLES
max_char = 0        # Tracks char ID for error testing.
max_guild = 0       # Tracks guild ID for error testing.
mfc = 0             # Tracks how many little files were counted.
write_data_loc = 'C:/Users/nrangara/Downloads/WorldOfWarcraft/output/'
              # Adjust to your dir as needed.
write_data_file = 'wowah_data.csv'
write_data_filename = write_data_loc + write_data_file
the_dir = 'C:/Users/nrangara/Downloads/WorldOfWarcraft/WoWAH/'                          # This i
s where the WoWAH folders are located, adjust as needed. Have them in their own subdir.

#REGEX
line_re = re.compile(r'^.*"[\d+],\s(.*),\s?(\d+),\s?(\d+),\s?(\d*),\s?(\d*),\s?([A-Z].*),\s?([A-
Z].*),\s?([A-Z].*)".*$')
#                           dummy   time1   seq2    3char   4guild    5level     6race     7charc
lass      8zone
#line_re = re.compile(r'^.*"[\d+],\s(.*),\s(\d+),\s?(\d+),\s?(\d*),\s?(\d*),\s?([A-Z][0-9].
*),\s?([A-Z][0-9].*),\s?([A-Z][0-9].*)".*$')
#                           dummy   time1   seq2  3char    4guild   5level       6race       7cha
rclass   8zone
#line_re = re.compile(r'^.*"[\d+],\s(.*),\s(\d+),\s?(\d+),\s?(\d*),.*".*$')
# REGEX NOTES
# groups: 1=timestamp, 3=avatarID, 4=guild.
# [1] = "0, 03/30/06 23:59:49, 1,10772, , 1, Orc, Warrior, Orgrimmar, , 0",
#       "0, 01/10/09 00:03:50, 1,55517, , 3, Orc, Warlock, Orgrimmar, WARLOCK, 0", -- [1]
#       "0, 01/10/09 00:04:10, 5,4002,1, 75, Orc, Hunter, Zul'Gurub, HUNTER, 0", -- [26]
#       "0, 01/10/09 00:04:10, 5,78122,342, 80, Orc, Hunter, The Storm Peaks, HUNTER, 0", -- [3
2]
#       "0, 01/10/09 00:08:04, 51,64635,161, 80, Blood Elf, Paladin, The Obsidian Sanctum, PALAD
IN, 0", -- [447]
# dummy, query time, query sequence number, avatar ID, guild, level, race, class, zone, dummy, d
ummy

def get_subdirs(the_folder):
    this_list = []
    this_list = os.listdir(the_folder)
    print ('From get_subdirs, a list is: ', this_list    )  # Printing for error control.
    for item in this_list:
        if item.startswith('.'):
            this_list.remove(item)
    return(this_list)
# End of get_subdirs
# '.DS_Store'

def get_file_list(the_folder):              # Yes these two are the same, just diff names.
    this_list = []
```

```
    this_list = os.listdir(the_folder)
    for item in this_list:
        if item.startswith('.'):
            this_list.remove(item)
    return(this_list)
# End of get_file_list

def parse_and_write(file, output_file):
    for line in file:                           # Oh the first "line" is a hard return???
#       print 'A line is: ', line
        data = line_re.match(line)
        #print (line)
        if data is not None:
            #print ("Matched the regex.")
            timestamp = data.group(1)
            char = data.group(3)
            level = data.group(5)
            race = data.group(6)
            charclass = data.group(7)
            zone = data.group(8)

            if data.group(4) is not(''):
                guild = data.group(4)
            else:
                guild = '-1'   # Note there are some missing values, i.e. errant -1.
            #print (timestamp)    # This is so you can keep track of where it is. Max Jan 2009 I
IRC.

            new_line = char + ',' + level + ',' + race + ',' + charclass + ',' + zone +  ',' + g
uild + ',' + timestamp + '\n'
            #new_line = char + ',' + guild + ',' + timestamp + '\n'
            output_file.write(new_line)

        #else:
            #print ("Didn't match the regex.")


# End of parse_and_write
# Note the two diff formats they use in the files, it changes partway through:
# [1] = "0, 03/30/06 23:59:49, 1,10772, , 1, Orc, Warrior, Orgrimmar, , 0",
#       "0, 01/10/09 00:03:50, 1,55517, , 3, Orc, Warlock, Orgrimmar, WARLOCK, 0", -- [1]
# dummy, query time, query sequence number, avatar ID, guild, level, race, class, zone, dummy, d
ummy

def read_tree(output_file):
    global the_dir
    months_folders = get_subdirs(the_dir)       # This is why the subdirs should be in their own
 location that you set in the vars section up top.
    for folder in months_folders:                                   # Run isdir(dir) first, try/
except. Make sure no funny folders/dirs.
        folder = the_dir + folder                                   # Expands the folder name to
 the long version.
        day_folders = get_subdirs(folder)
        for day_folder in day_folders:
```

```
            day_folder = folder + '/' + day_folder
            file_list = get_file_list(day_folder)
            for file in file_list:
                try:
                    file = day_folder + '/' + file
                    with open(file, 'r',encoding='utf8') as f:
                        this_file = f.readlines()                          # Should read the who
le file as a string?
                        parse_and_write(this_file, output_file)
                except IOError:
                    print ('Error opening hoped for data-text file,', str(file), ', reason: ', I
OError)
# End of read_tree

def main():
    #open write file here
    output_file = open(write_data_filename, 'a',encoding='utf8')    # 'a' is very important, it
 appends the new data to the big file.
    fieldnames = ('char, level, race, charclass, zone, dummy1, dummy2, guild, timestamp\n')
    output_file.write(fieldnames)
    start_time = time.time()
    read_tree(output_file)
    #close write file here
    output_file.close()
    spent_time = time.time() - start_time
    mins_spent = int(spent_time / 60)
    secs_remainder = int(spent_time % 60)
    print ('Time of process: ', mins_spent, ':', secs_remainder)     # 13m:42s on iMac. Also 14
m:39s another time.

#    print 'Files scanned (or tried), ', mfc     # 138,084
#    print 'Max Chars: ', max_char               # They say 91,065 ">= 1" but it starts at 0, my
 count says: 91064 + 1 = 91,065.
#    print 'Max Guilds: ', max_guild             # They say "An integer within [1, 513]" but no
 since they start at 0. 512 + 1 = 513.
# End of main


# Main call

main()
```

```
## From get_subdirs, a list is:  ['2006_01_03', '2006_04_06', '2006_07_09', '2006_10_12', '2007_
01_03', '2007_04_06', '2007_07_09', '2007_10_12', '2008_01_03', '2008_04_06', '2008_07_09', '200
8_10_12', '2009_01']
## From get_subdirs, a list is:  ['2006-01-01', '2006-01-02', '2006-01-03', '2006-01-04', '2006-
01-05', '2006-01-06', '2006-01-07', '2006-01-08', '2006-01-09', '2006-01-10', '2006-01-11', '200
6-01-12', '2006-01-13', '2006-01-14', '2006-01-15', '2006-01-16', '2006-01-17', '2006-01-18', '2
006-01-19', '2006-01-20', '2006-01-21', '2006-01-22', '2006-01-23', '2006-01-24', '2006-01-25',
'2006-01-26', '2006-01-27', '2006-01-28', '2006-01-29', '2006-01-30', '2006-01-31', '2006-02-0
1', '2006-02-02', '2006-02-03', '2006-02-04', '2006-02-05', '2006-02-06', '2006-02-07', '2006-02
-08', '2006-02-09', '2006-02-10', '2006-02-11', '2006-02-12', '2006-02-13', '2006-02-14', '2006-
02-15', '2006-02-16', '2006-02-17', '2006-02-18', '2006-02-19', '2006-02-20', '2006-02-21', '200
6-02-22', '2006-02-23', '2006-02-24', '2006-02-25', '2006-02-26', '2006-02-27', '2006-02-28', '2
006-03-01', '2006-03-02', '2006-03-03', '2006-03-04', '2006-03-05', '2006-03-06', '2006-03-07',
'2006-03-08', '2006-03-09', '2006-03-10', '2006-03-11', '2006-03-12', '2006-03-13', '2006-03-1
4', '2006-03-15', '2006-03-16', '2006-03-17', '2006-03-18', '2006-03-19', '2006-03-20', '2006-03
-21', '2006-03-22', '2006-03-23', '2006-03-24', '2006-03-25', '2006-03-26', '2006-03-27', '2006-
03-28', '2006-03-29', '2006-03-30', '2006-03-31']
## From get_subdirs, a list is:  ['2006-04-01', '2006-04-02', '2006-04-03', '2006-04-04', '2006-
04-05', '2006-04-06', '2006-04-07', '2006-04-08', '2006-04-09', '2006-04-10', '2006-04-11', '200
6-04-12', '2006-04-13', '2006-04-14', '2006-04-15', '2006-04-16', '2006-04-17', '2006-04-18', '2
006-04-19', '2006-04-20', '2006-04-21', '2006-04-23', '2006-04-24', '2006-04-25', '2006-04-26',
'2006-04-27', '2006-04-28', '2006-04-29', '2006-04-30', '2006-05-01', '2006-05-02', '2006-05-0
3', '2006-05-04', '2006-05-05', '2006-05-06', '2006-05-07', '2006-05-08', '2006-05-09', '2006-05
-10', '2006-05-11', '2006-05-12', '2006-05-13', '2006-05-14', '2006-05-15', '2006-05-16', '2006-
05-17', '2006-05-18', '2006-05-19', '2006-05-20', '2006-05-21', '2006-05-22', '2006-05-23', '200
6-05-24', '2006-05-25', '2006-05-26', '2006-05-27', '2006-05-28', '2006-05-29', '2006-05-30', '2
006-05-31', '2006-06-01', '2006-06-02', '2006-06-03', '2006-06-04', '2006-06-05', '2006-06-06',
'2006-06-07', '2006-06-08', '2006-06-09', '2006-06-10', '2006-06-11', '2006-06-12', '2006-06-1
3', '2006-06-14', '2006-06-15', '2006-06-16', '2006-06-17', '2006-06-18', '2006-06-19', '2006-06
-20', '2006-06-21', '2006-06-22', '2006-06-23', '2006-06-24', '2006-06-25', '2006-06-26', '2006-
06-27', '2006-06-28', '2006-06-29', '2006-06-30']
## From get_subdirs, a list is:  ['2006-07-01', '2006-07-02', '2006-07-03', '2006-07-04', '2006-
07-05', '2006-07-06', '2006-07-07', '2006-07-08', '2006-07-09', '2006-07-10', '2006-07-11', '200
6-07-12', '2006-07-13', '2006-07-14', '2006-07-15', '2006-07-16', '2006-07-17', '2006-07-18', '2
006-07-19', '2006-07-20', '2006-07-21', '2006-07-22', '2006-07-23', '2006-07-24', '2006-07-25',
'2006-07-26', '2006-07-27', '2006-07-28', '2006-07-29', '2006-07-30', '2006-07-31', '2006-08-0
1', '2006-08-02', '2006-08-03', '2006-08-04', '2006-08-05', '2006-08-06', '2006-08-07', '2006-08
-08', '2006-08-09', '2006-08-10', '2006-08-11', '2006-08-12', '2006-08-13', '2006-08-14', '2006-
08-15', '2006-08-16', '2006-08-17', '2006-08-18', '2006-08-19', '2006-08-20', '2006-08-21', '200
6-08-22', '2006-08-23', '2006-08-24', '2006-08-25', '2006-08-26', '2006-08-27', '2006-08-28', '2
006-08-29', '2006-08-30', '2006-08-31', '2006-09-01', '2006-09-02', '2006-09-03', '2006-09-04',
'2006-09-05', '2006-09-06', '2006-09-07', '2006-09-08', '2006-09-09', '2006-09-10', '2006-09-1
1', '2006-09-12', '2006-09-13', '2006-09-14', '2006-09-15', '2006-09-16', '2006-09-17', '2006-09
-18', '2006-09-19', '2006-09-20', '2006-09-21', '2006-09-22', '2006-09-23', '2006-09-24', '2006-
09-25', '2006-09-26', '2006-09-27', '2006-09-28', '2006-09-29', '2006-09-30']
## From get_subdirs, a list is:  ['2006-10-01', '2006-10-02', '2006-10-03', '2006-10-04', '2006-
10-05', '2006-10-06', '2006-10-07', '2006-10-08', '2006-10-09', '2006-10-10', '2006-10-11', '200
6-10-12', '2006-10-13', '2006-10-14', '2006-10-15', '2006-10-16', '2006-10-17', '2006-10-18', '2
006-10-19', '2006-10-20', '2006-10-21', '2006-10-22', '2006-10-23', '2006-10-24', '2006-10-25',
'2006-10-26', '2006-10-27', '2006-10-28', '2006-10-29', '2006-10-30', '2006-10-31', '2006-11-0
1', '2006-11-02', '2006-11-03', '2006-11-04', '2006-11-05', '2006-11-06', '2006-11-07', '2006-11
-08', '2006-11-09', '2006-11-10', '2006-11-11', '2006-11-12', '2006-11-13', '2006-11-14', '2006-
11-15', '2006-11-16', '2006-11-17', '2006-11-18', '2006-11-19', '2006-11-20', '2006-11-21', '200
```

```
6-11-22', '2006-11-23', '2006-11-24', '2006-11-25', '2006-11-26', '2006-11-27', '2006-11-28', '2
006-11-29', '2006-11-30', '2006-12-01', '2006-12-02', '2006-12-03', '2006-12-04', '2006-12-05',
'2006-12-06', '2006-12-07', '2006-12-08', '2006-12-09', '2006-12-10', '2006-12-11', '2006-12-1
2', '2006-12-13', '2006-12-14', '2006-12-15', '2006-12-16', '2006-12-17', '2006-12-18', '2006-12
-19', '2006-12-20', '2006-12-21', '2006-12-22', '2006-12-23', '2006-12-24', '2006-12-25', '2006-
12-26', '2006-12-27', '2006-12-28', '2006-12-29', '2006-12-30', '2006-12-31']
## From get_subdirs, a list is:  ['2007-01-01', '2007-01-02', '2007-01-03', '2007-01-04', '2007-
01-05', '2007-01-06', '2007-01-07', '2007-01-08', '2007-01-09', '2007-01-10', '2007-01-11', '200
7-01-12', '2007-01-13', '2007-01-14', '2007-01-15', '2007-01-16', '2007-01-17', '2007-01-18', '2
007-01-19', '2007-01-20', '2007-01-22', '2007-01-23', '2007-01-24', '2007-01-25', '2007-01-26',
'2007-01-27', '2007-01-28', '2007-01-29', '2007-01-30', '2007-01-31', '2007-02-01', '2007-02-0
2', '2007-02-03', '2007-02-04', '2007-02-05', '2007-02-06', '2007-02-07', '2007-02-08', '2007-02
-09', '2007-02-10', '2007-02-11', '2007-02-12', '2007-02-13', '2007-02-14', '2007-02-15', '2007-
02-16', '2007-02-17', '2007-02-18', '2007-02-19', '2007-02-20', '2007-02-21', '2007-02-22', '200
7-02-23', '2007-02-24', '2007-02-25', '2007-02-26', '2007-02-27', '2007-02-28', '2007-03-01', '2
007-03-02', '2007-03-03', '2007-03-04', '2007-03-05', '2007-03-06', '2007-03-07', '2007-03-08',
'2007-03-09', '2007-03-10', '2007-03-11', '2007-03-12', '2007-03-13', '2007-03-14', '2007-03-1
5', '2007-03-16', '2007-03-17', '2007-03-18', '2007-03-19', '2007-03-20', '2007-03-24', '2007-03
-25', '2007-03-26', '2007-03-27', '2007-03-28', '2007-03-29', '2007-03-30', '2007-03-31']
## From get_subdirs, a list is:  ['2007-04-01', '2007-04-02', '2007-04-03', '2007-04-04', '2007-
04-05', '2007-04-06', '2007-04-07', '2007-04-08', '2007-04-09', '2007-04-10', '2007-04-11', '200
7-04-12', '2007-04-13', '2007-04-14', '2007-04-15', '2007-04-16', '2007-04-17', '2007-04-18', '2
007-04-19', '2007-04-20', '2007-04-21', '2007-04-22', '2007-04-23', '2007-04-24', '2007-04-25',
'2007-04-26', '2007-04-27', '2007-04-28', '2007-04-29', '2007-04-30', '2007-05-01', '2007-05-0
2', '2007-05-03', '2007-05-04', '2007-05-05', '2007-05-06', '2007-05-07', '2007-05-08', '2007-05
-09', '2007-05-10', '2007-05-11', '2007-05-12', '2007-05-13', '2007-05-14', '2007-05-15', '2007-
05-16', '2007-05-17', '2007-05-18', '2007-05-19', '2007-05-20', '2007-05-21', '2007-05-22', '200
7-05-23', '2007-05-24', '2007-05-25', '2007-05-31', '2007-06-01', '2007-06-02', '2007-06-03', '2
007-06-04', '2007-06-05', '2007-06-06', '2007-06-07', '2007-06-08', '2007-06-09', '2007-06-10',
'2007-06-11', '2007-06-12', '2007-06-13', '2007-06-14', '2007-06-15', '2007-06-16', '2007-06-1
7', '2007-06-18', '2007-06-19', '2007-06-20', '2007-06-21', '2007-06-22', '2007-06-23', '2007-06
-24', '2007-06-25', '2007-06-26', '2007-06-27', '2007-06-28', '2007-06-29', '2007-06-30']
## From get_subdirs, a list is:  ['2007-07-01', '2007-07-02', '2007-07-03', '2007-07-04', '2007-
07-05', '2007-07-06', '2007-07-07', '2007-07-08', '2007-07-09', '2007-07-10', '2007-07-11', '200
7-07-12', '2007-07-13', '2007-07-14', '2007-07-15', '2007-07-16', '2007-07-17', '2007-07-18', '2
007-07-19', '2007-07-20', '2007-07-21', '2007-07-22', '2007-07-23', '2007-07-24', '2007-07-25',
'2007-07-26', '2007-07-27', '2007-07-28', '2007-07-29', '2007-07-30', '2007-07-31', '2007-08-0
1', '2007-08-02', '2007-08-03', '2007-08-04', '2007-08-05', '2007-08-06', '2007-08-07', '2007-08
-08', '2007-08-09', '2007-08-10', '2007-08-11', '2007-08-12', '2007-08-13', '2007-08-14', '2007-
08-15', '2007-08-16', '2007-08-17', '2007-08-18', '2007-08-19', '2007-08-20', '2007-08-21', '200
7-08-22', '2007-08-23', '2007-08-24', '2007-08-25', '2007-08-26', '2007-08-27', '2007-08-28', '2
007-08-29', '2007-08-30', '2007-08-31', '2007-09-01', '2007-09-02', '2007-09-03', '2007-09-04',
'2007-09-05', '2007-09-06', '2007-09-07', '2007-09-08', '2007-09-09', '2007-09-10', '2007-09-1
1', '2007-09-12', '2007-09-13', '2007-09-14', '2007-09-15', '2007-09-16', '2007-09-17', '2007-09
-18', '2007-09-19', '2007-09-20', '2007-09-21', '2007-09-22', '2007-09-23', '2007-09-24', '2007-
09-25', '2007-09-26', '2007-09-27', '2007-09-28', '2007-09-29', '2007-09-30']
## From get_subdirs, a list is:  ['2007-10-01', '2007-10-02', '2007-10-03', '2007-10-04', '2007-
10-05', '2007-10-06', '2007-10-07', '2007-10-08', '2007-10-09', '2007-10-10', '2007-10-11', '200
7-10-12', '2007-10-13', '2007-10-14', '2007-10-15', '2007-10-16', '2007-10-17', '2007-10-18', '2
007-10-19', '2007-10-20', '2007-10-21', '2007-10-22', '2007-10-23', '2007-10-24', '2007-10-25',
'2007-10-26', '2007-10-27', '2007-10-28', '2007-10-29', '2007-10-30', '2007-10-31', '2007-11-0
1', '2007-11-02', '2007-11-03', '2007-11-04', '2007-11-05', '2007-11-06', '2007-11-07', '2007-11
-08', '2007-11-09', '2007-11-10', '2007-11-11', '2007-11-12', '2007-11-13', '2007-11-14', '2007-
11-15', '2007-11-16', '2007-11-17', '2007-11-18', '2007-11-19', '2007-11-20', '2007-11-21', '200
```

```
7-11-22', '2007-11-23', '2007-11-24', '2007-11-25', '2007-11-26', '2007-11-27', '2007-11-28', '2
007-11-29', '2007-11-30', '2007-12-01', '2007-12-02', '2007-12-03', '2007-12-04', '2007-12-05',
'2007-12-06', '2007-12-07', '2007-12-08', '2007-12-09', '2007-12-10', '2007-12-11', '2007-12-1
2', '2007-12-13', '2007-12-14', '2007-12-15', '2007-12-16', '2007-12-17', '2007-12-18', '2007-12
-19', '2007-12-20', '2007-12-21', '2007-12-22', '2007-12-23', '2007-12-24', '2007-12-25', '2007-
12-26', '2007-12-27', '2007-12-28', '2007-12-29', '2007-12-30', '2007-12-31']
## From get_subdirs, a list is:  ['2008-01-01', '2008-01-02', '2008-01-03', '2008-01-04', '2008-
01-05', '2008-01-06', '2008-01-07', '2008-01-08', '2008-01-09', '2008-01-10', '2008-01-11', '200
8-01-12', '2008-01-13', '2008-01-14', '2008-01-15', '2008-01-16', '2008-01-17', '2008-01-18', '2
008-01-19', '2008-01-20', '2008-01-21', '2008-01-22', '2008-01-23', '2008-01-24', '2008-01-25',
'2008-01-26', '2008-01-27', '2008-01-28', '2008-01-29', '2008-01-30', '2008-01-31', '2008-02-0
1', '2008-02-02', '2008-02-03', '2008-02-04', '2008-02-05', '2008-02-06', '2008-02-07', '2008-02
-08', '2008-02-09', '2008-02-10', '2008-02-11', '2008-02-12', '2008-02-13', '2008-02-14', '2008-
02-15', '2008-02-16', '2008-02-17', '2008-02-18', '2008-02-19', '2008-02-20', '2008-02-21', '200
8-02-22', '2008-02-23', '2008-02-24', '2008-02-25', '2008-02-26', '2008-02-27', '2008-02-28', '2
008-02-29', '2008-03-01', '2008-03-02', '2008-03-03', '2008-03-04', '2008-03-05', '2008-03-06',
'2008-03-07', '2008-03-08', '2008-03-09', '2008-03-10', '2008-03-11', '2008-03-12', '2008-03-1
3', '2008-03-14', '2008-03-15', '2008-03-16', '2008-03-17', '2008-03-18', '2008-03-19', '2008-03
-20', '2008-03-21', '2008-03-22', '2008-03-23', '2008-03-24', '2008-03-25', '2008-03-26', '2008-
03-27', '2008-03-28', '2008-03-29', '2008-03-30', '2008-03-31']
## From get_subdirs, a list is:  ['2008-04-01', '2008-04-02', '2008-04-03', '2008-04-07', '2008-
04-08', '2008-04-09', '2008-04-10', '2008-04-11', '2008-04-12', '2008-04-13', '2008-04-14', '200
8-04-15', '2008-04-16', '2008-04-17', '2008-04-18', '2008-04-19', '2008-04-20', '2008-04-21', '2
008-04-22', '2008-04-23', '2008-04-24', '2008-04-25', '2008-04-26', '2008-04-27', '2008-04-28',
'2008-04-29', '2008-04-30', '2008-05-06', '2008-05-07', '2008-05-08', '2008-05-09', '2008-05-1
0', '2008-05-11', '2008-05-12', '2008-05-13', '2008-05-14', '2008-05-15', '2008-05-16', '2008-05
-19', '2008-05-20', '2008-05-21', '2008-05-22', '2008-05-23', '2008-05-24', '2008-05-25', '2008-
05-26', '2008-05-27', '2008-05-28', '2008-05-29', '2008-05-30', '2008-05-31', '2008-06-01', '200
8-06-02', '2008-06-03', '2008-06-04', '2008-06-05', '2008-06-06', '2008-06-07', '2008-06-08', '2
008-06-09', '2008-06-10', '2008-06-11', '2008-06-12', '2008-06-13', '2008-06-18', '2008-06-19',
'2008-06-20', '2008-06-21', '2008-06-22', '2008-06-23', '2008-06-24', '2008-06-25', '2008-06-2
6', '2008-06-27', '2008-06-28', '2008-06-29', '2008-06-30']
## From get_subdirs, a list is:  ['2008-07-01', '2008-07-02', '2008-07-03', '2008-07-04', '2008-
07-05', '2008-07-06', '2008-07-07', '2008-07-08', '2008-07-09', '2008-07-10', '2008-07-11', '200
8-07-12', '2008-07-13', '2008-07-14', '2008-07-15', '2008-07-16', '2008-07-17', '2008-07-18', '2
008-07-19', '2008-07-20', '2008-07-21', '2008-07-22', '2008-07-23', '2008-07-24', '2008-07-25',
'2008-07-26', '2008-07-27', '2008-07-28', '2008-07-29', '2008-07-30', '2008-07-31', '2008-08-0
1', '2008-08-02', '2008-08-03', '2008-08-04', '2008-08-05', '2008-08-06', '2008-08-07', '2008-08
-08', '2008-08-09', '2008-08-10', '2008-08-11', '2008-08-12', '2008-08-13', '2008-08-14', '2008-
08-15', '2008-08-16', '2008-08-17', '2008-08-18', '2008-08-19', '2008-08-20', '2008-08-21', '200
8-08-22', '2008-08-23', '2008-08-24', '2008-08-25', '2008-08-26', '2008-08-27', '2008-08-28', '2
008-08-29', '2008-08-30', '2008-08-31', '2008-09-01', '2008-09-02', '2008-09-03', '2008-09-04',
'2008-09-05', '2008-09-06', '2008-09-07', '2008-09-08', '2008-09-09', '2008-09-10', '2008-09-1
1', '2008-09-12', '2008-09-13', '2008-09-14', '2008-09-15', '2008-09-16', '2008-09-17', '2008-09
-18', '2008-09-19', '2008-09-20', '2008-09-21', '2008-09-22', '2008-09-23', '2008-09-24', '2008-
09-25', '2008-09-26', '2008-09-27', '2008-09-28', '2008-09-29', '2008-09-30']
## From get_subdirs, a list is:  ['2008-10-01', '2008-10-02', '2008-10-03', '2008-10-04', '2008-
10-05', '2008-10-06', '2008-10-07', '2008-10-08', '2008-10-09', '2008-10-10', '2008-10-11', '200
8-10-12', '2008-10-13', '2008-10-14', '2008-10-15', '2008-10-16', '2008-10-17', '2008-10-18', '2
008-10-19', '2008-10-20', '2008-10-21', '2008-10-22', '2008-10-23', '2008-10-24', '2008-10-25',
'2008-10-26', '2008-10-27', '2008-10-28', '2008-10-29', '2008-10-30', '2008-10-31', '2008-11-0
1', '2008-11-02', '2008-11-03', '2008-11-04', '2008-11-05', '2008-11-06', '2008-11-07', '2008-11
-08', '2008-11-09', '2008-11-10', '2008-11-11', '2008-11-12', '2008-11-13', '2008-11-14', '2008-
11-15', '2008-11-16', '2008-11-17', '2008-11-18', '2008-11-19', '2008-11-20', '2008-11-21', '200
```

```
8-11-22', '2008-11-23', '2008-11-24', '2008-11-25', '2008-11-26', '2008-11-27', '2008-11-28', '2
008-11-29', '2008-11-30', '2008-12-01', '2008-12-02', '2008-12-03', '2008-12-04', '2008-12-05',
'2008-12-06', '2008-12-07', '2008-12-08', '2008-12-09', '2008-12-10', '2008-12-11', '2008-12-1
2', '2008-12-13', '2008-12-14', '2008-12-15', '2008-12-16', '2008-12-17', '2008-12-18', '2008-12
-19', '2008-12-20', '2008-12-21', '2008-12-22', '2008-12-23', '2008-12-24', '2008-12-25', '2008-
12-26', '2008-12-27', '2008-12-28', '2008-12-29', '2008-12-30', '2008-12-31']
## From get_subdirs, a list is:  ['2009-01-01', '2009-01-02', '2009-01-03', '2009-01-04', '2009-
01-05', '2009-01-06', '2009-01-07', '2009-01-08', '2009-01-09', '2009-01-10']
## Time of process:  8 : 55
```