

Homework 5: Processing Structured Data

Due Wednesday, October 29, 3p (*note: changed to start of class.*)

- hw5-exercises.py program
- hw5-part2.py program
- hw5_output.csv output from your program
- an image of your Google Chart

Part 0: Get the homework files

You'll need the homework files. Get them by typing the following lines in terminal:

```
cd ~  
./getHW.sh 5
```

Don't forget that if you already have Eclipse open, you may need to right click on the homeworks project and "refresh" it, to make the new folder 5 appear.

Part 1: Exercises

Open `hw5-exercises.py` and follow instructions in there.

Part 2: Processing the feed

Open `hw5-part2.py`. It contains code with everything you need to start reading in and processing data from the HCDE 310 Facebook group. It also contains instructions on what you should be doing throughout the assignment. Any lines beginning with `#<your-code-here>` indicate that there is an exercise or task that you need to complete.

2.1 Overview

In this assignment, you will be processing real data from Facebook. It's in a format that is generally convenient to work with – but it still requires some extra work to get only the parts you need for the main tasks in this assignment. We will give you code to read in the JSON-formatted Facebook data as a list of dictionaries. Each dictionary represents a post to the group. The Building Block exercises will guide you through developing components that will extract the necessary parts of the post dictionaries used in the main tasks. You will reference the information below as you follow along in `hw5-part2.py`.

2.2 Building Blocks

There are three building blocks used in your tasks:

1. Extracting the shortened name of the author (contributor) of the post

```
-----building block 1-----
{
  "id": "860888250596380",
  "name": "Lisa Hu"
}
Lisa Hu
Lisa H.
Sabrina W.
```

2. Extracting the number of comments a post received

```
Comment count: 3
3
0
```

3. Extracting the number of likes a post received

```
Likes: 2
0
2
```

hw5-part2.py will guide you through writing these three functions.

2.3 Tasks

1. *Printing shortened names, number of comments and likes for each post*

For each post in the posts list, you should print out a single line of the form: 'Bob H. posted to the group and received 5 comments and 2 likes.'

It is not necessary to use the singular form 'comments' or 'likes' when the count is equal to one (as you did in the previous week). The first few lines of output should look like:

```
-----Task 1-----
Nitaya M. posted to the group and received 1 comments and 0 likes
Samuel M. posted to the group and received 9 comments and 0 likes
Sabrina W. posted to the group and received 3 comments and 0 likes
Kendall A. posted to the group and received 2 comments and 1 likes
Keting C. posted to the group and received 4 comments and 0 likes
Chia-Fang C. posted to the group and received 0 comments and 3 likes
Lisa H. posted to the group and received 1 comments and 6 likes
Xiaochen Y. posted to the group and received 4 comments and 0 likes
```

2. Counting the number of posts for each user and the total number of likes and comments they received

Follow a procedure similar to the past two assignments using building blocks (1-3) to count the number of posts (“contributor counts”) for each user, as well as the total number of comments and posts they received. Using each user’s shortened name as the key, you will need to use three dictionaries, `post_counts`, `like_counts`, and `comment_counts` to keep a running total of how many times the user posted, and the total number of likes and comments the user received.

Output should look like (order may vary!):

-----Task 2-----

```
Nicole T. 2 0 4
Samuel M. 2 0 11
Nitaya M. 2 0 3
Sabrina W. 2 0 6
Jessica B. 2 0 4
Sean M. 12 23 12
Chia-Fang C. 4 8 4
Sami W. 4 0 23
Lisa H. 2 6 2
Stephanie G. 1 2 3
Keting C. 1 0 4
Max S. 1 0 1
Anne Z. 1 2 3
Yoanna D. 2 8 5
Candelario P. 1 0 6
Tristan S. 1 0 1
Ashley D. 2 3 2
Wei K. 1 0 6
Kyle N. 2 0 4
Samuel S. 3 5 5
Danny C. 1 0 0
Kendall A. 2 1 4
Autumn G. 2 4 2
Anu M. 1 0 9
Perry M. 3 2 20
Xiaochen Y. 2 0 11
```

3. Writing contributor counts to a CSV file

Write out your data to a file named `hw5_output.csv`.

The first line should say:

```
name,posts,likes,comments
```

Then, you should print out the following information for every user that had appeared in posts:

shortname,num_posts,total_likes,total_comments

e.g. if Bob H. made 4 posts, received 6 comments in total, and 2 likes in total, you should have a corresponding line in your file that says:

Bob H.,4,2,6

Here is what the first few lines of hw5_output.csv should look like (note: the output following the first line may be in a different order on your computer):

```
name,posts,likes,comments
Nicole T.,2,0,4
Samuel M.,2,0,11
Nitaya M.,2,0,3
Sabrina W.,2,0,6
Jessica B.,2,0,4
Sean M.,12,23,12
Chia-Fang C.,4,8,4
Sami W.,4,0,23
Lisa H.,2,6,2
Stephanie G.,1,2,3
Keting C.,1,0,4
Max S.,1,0,1
Anne Z.,1,2,3
Yoanna D.,2,8,5
Candelario P.,1,0,6
Tristan S.,1,0,1
Ashley D.,2,3,2
Wei K.,1,0,6
Kyle N.,2,0,4
Samuel S.,3,5,5
Danny C.,1,0,0
Kendall A.,2,1,4
Autumn G.,2,4,2
Anu M.,1,0,9
Perry M.,3,2,20
Xiaochen Y.,2,0,11
```

4. Creating a chart with Google Docs

You should be great at this by now. Follow last week's instructions to make a chart of this week's output. Again, you are encouraged to post your chart to Facebook.