**Instructions on how to compile and produce the desired output:**

I used PYTHON to complete this assignment. Firstly, to gear things up, I downloaded Anaconda and started coding on Spyder. It might come into surprise, but this is the first time I used Spyder or Anaconda. After exploring Anaconda, I preferred to go with Spyder.

The program runs on Spyder without any bugs. All one needs to do is to compile the program using one of the IDE’s. In addition, I have used Python 2.7. Installation of Anaconda version 2- 5.0.1, and use of Spyder sums up my description on compiling and running the code.

**Challenges faced during the assignment:**

Having no prior knowledge on data science, just a few ideas and some online tutorials on mind, I was putting myself on a challenge for this semester. After some classes, the topics were getting interesting and my interest towards data science and mining was growing. With the assignment 1 bringing all the topics that were covered, it made me research a lot and work on the basic ideas firstly and then to some deep learnings on python and data science.

One of the infuriating thing that happened to me was getting the IP blocked occasionally due to testing the code often. While I found about the sleep function to avoid the robotic behavior, I pretty much had a hard time implementing it and I ended up not using it. The assignment made me revise all the functions that python uses, as it had become a long time I had coded in Python after the substantial use of Java in the previous classes.

Other challenges were to tidy the data in the part 3 of the assignment. Extraction of data from repetitive tags within the same class, and parent tags made things more interesting.

I couldn’t compile part3 due to the ip blockage, but I will try my best and will send the tsv file soon.