

HW0: Text Processing in Python

Due: January 18th, 2018 (before end of the day)

In this assignment you will get familiar with Python and the NLTK toolkit for processing text. Your assignment is to work through the examples in Chapter 1 of the book located at <http://www.nltk.org/book/ch01.html>. We recommend that you follow all the exercises in there but at a minimum you need to complete the following ones:

1. Exercise 1.4
2. Exercises 3.1-3.4
3. Exercise 4.1
4. Exercise 8.28

Deliverables

Copy/paste screen shots of the output for exercises listed above in a document and submit as assignment 0.

How to Submit

You will need to do online submission of this. Submit your document or zipped file using this file request link <https://www.dropbox.com/request/xNzrBxgU1ETJT15mUkoW>. Name your file following this convention: `firstname.lastname.hw0`.

Software Requirements

You will need to install Python and NLTK on your machine. If you don't have them installed already, you can find a lot of help on how to do this online.

- <http://docs.python-guide.org/en/latest/intro/learning/>
- <https://www.kaggle.com/c/word2vec-nlp-tutorial/details/part-1-for-beginners-bag-of-words>
- <https://www.youtube.com/watch?v=FLZvOKSckxY>