NLP Notes

* Tokenization:- grouping parts of text together. Two types:- word tokenizers and sentence tokenizers.
* Lexicon:- words and their meanings. Kinda like a dictionary.
* Corpora:- body of text.