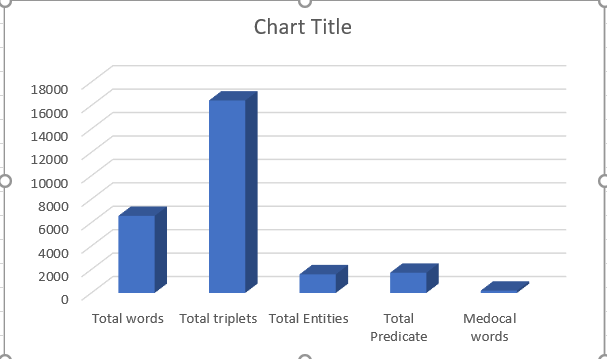
Name: Trinadha Raji Muppala

Class: 7

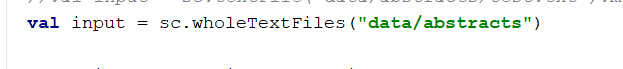
KDM Lab3

**Statistics for Parkinson's** **disease abstracts**



Finding All Statics:

1. Read all the abstracts:

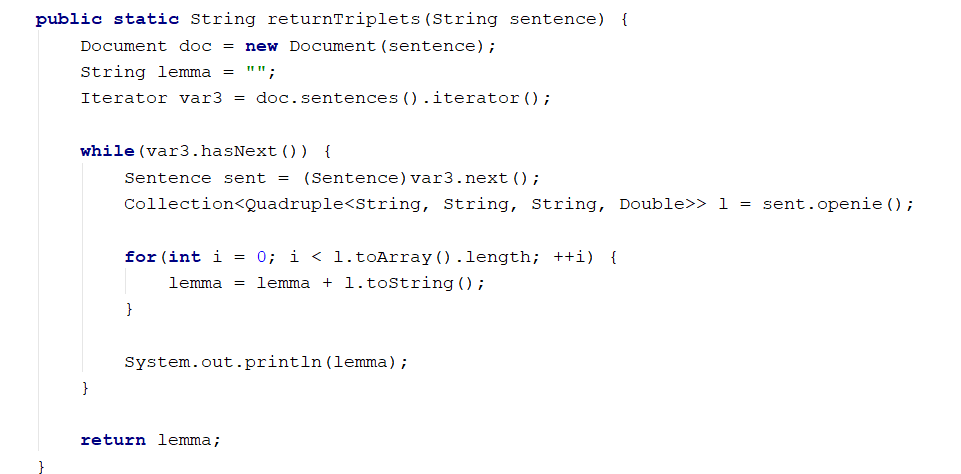


1. Find distinct words from abstract and find count

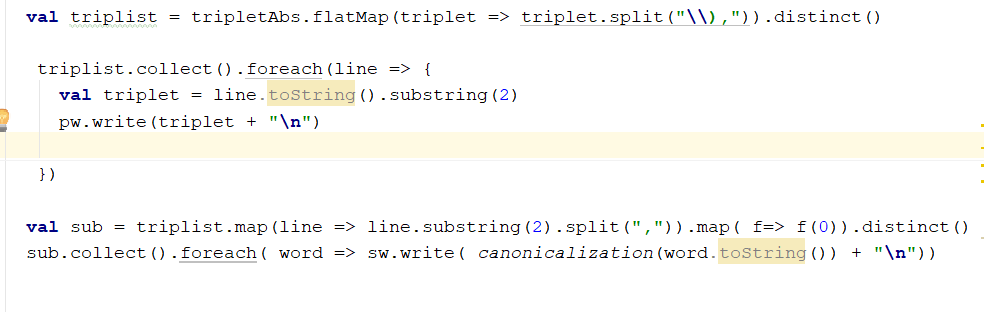




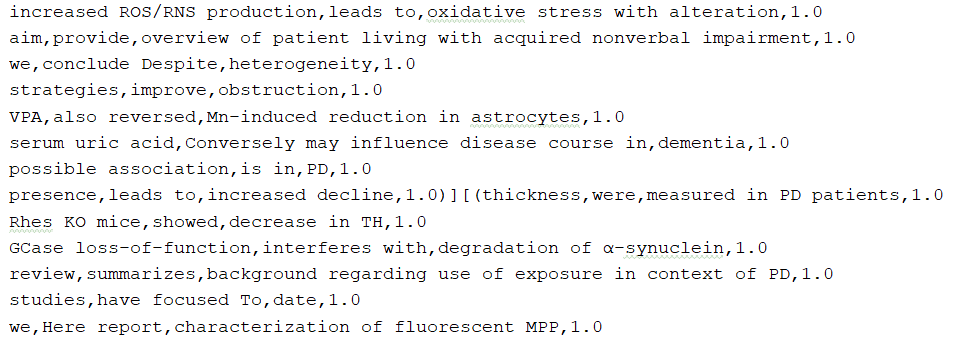
1. Find triplets using CORENLP



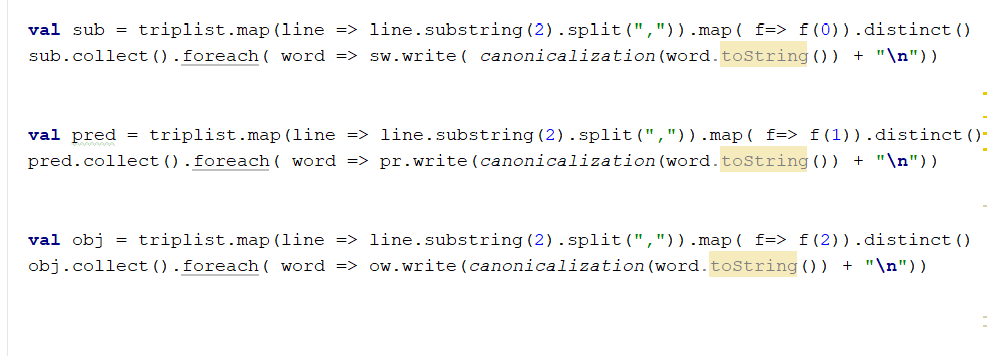
1. All the triplets are concatenated to one string. Split triplets using “),” and find out the count



Triplets:

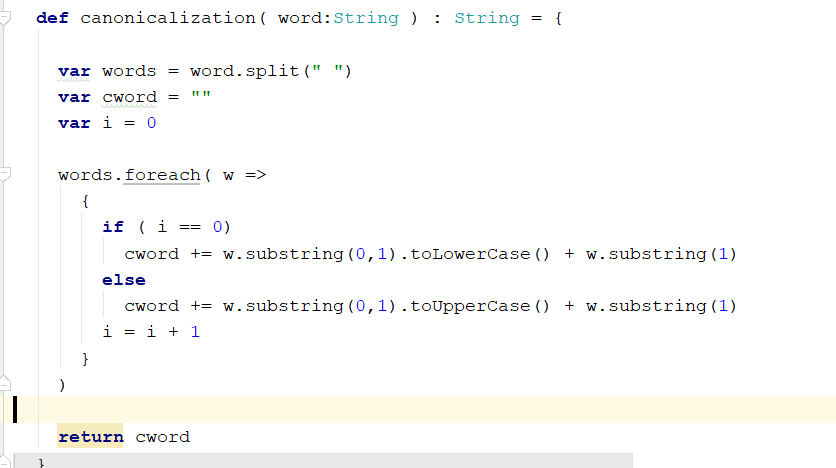


1. From the triplets find out subjects and objects, predicates splitting by “,” first second third words of the triplets

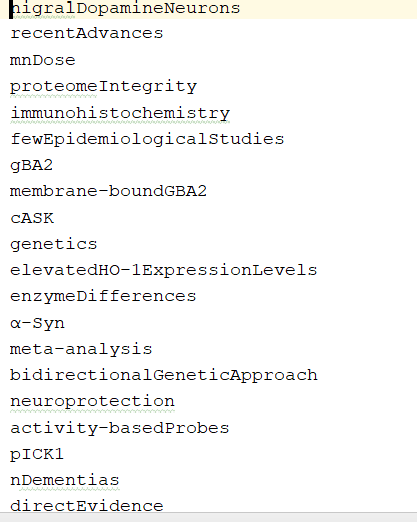


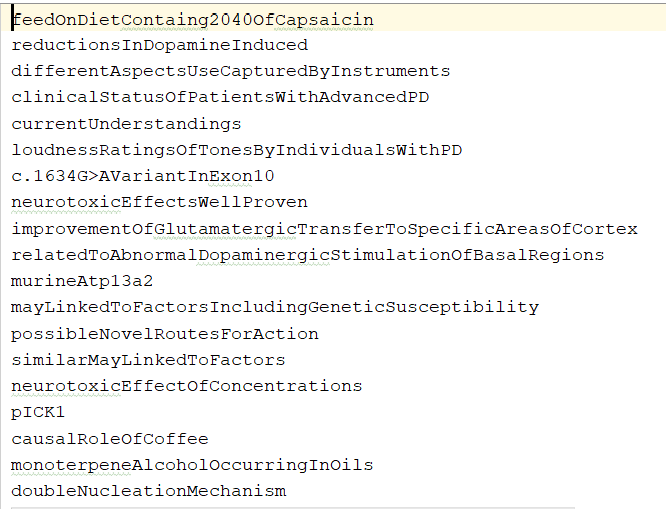
1. Change subject object predicate to canonical form

For first word change the first letter to lower, for rest of the words first letter to capital append rest of the word

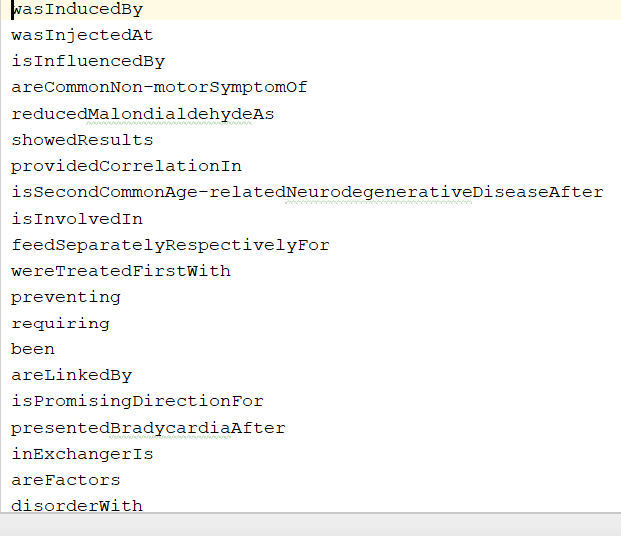


1. Entities





1. Predicate

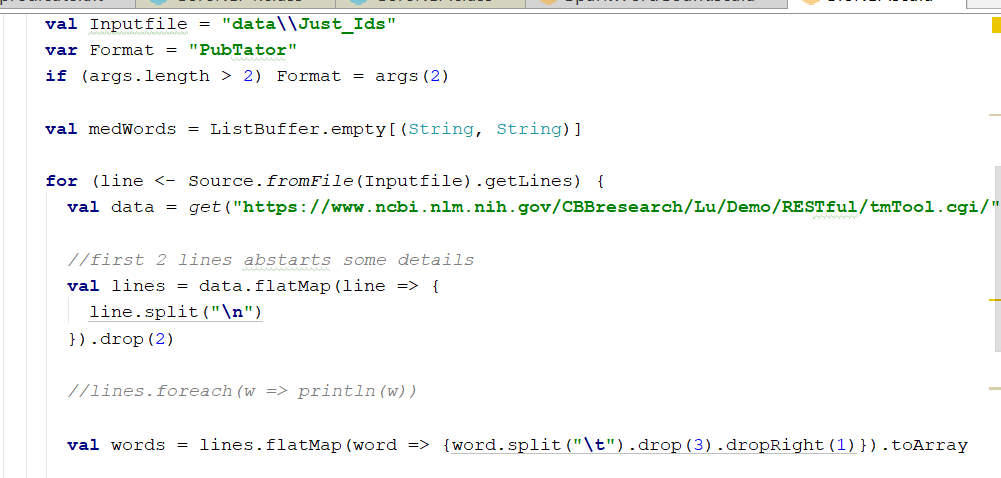


1. Medical words

Use the Ids from abstracts.

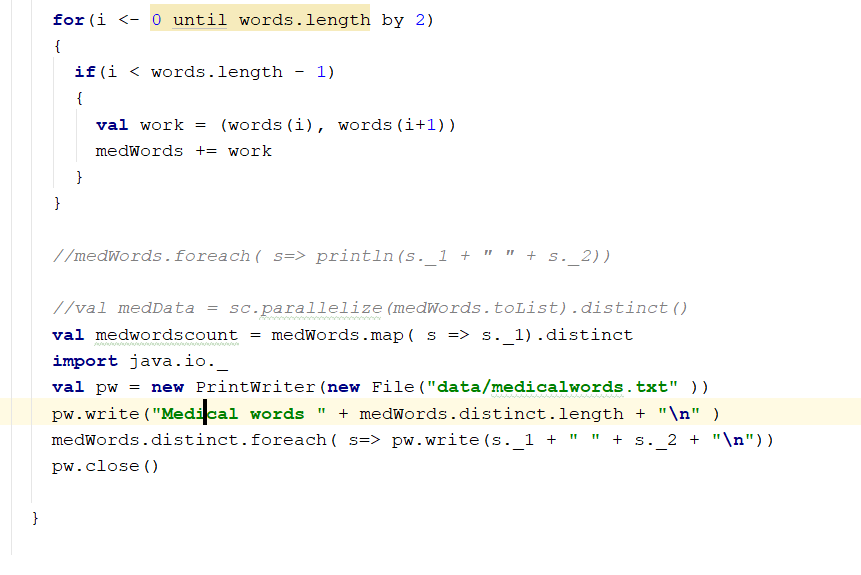
Call BIONLP api

Results will got tab delimited data which can provide Gene, Disease, Spices, Chemical. Drop first 3 and last word from the results



From the medical words, get distinct words write to medicalwords.txt file . Find out medical

words count



Medical words count, and medical words

