

Assignment #1

Instructor: Mayank Singh

Important Instructions: Use the “Tweets Dataset” that is attached along with this to perform the following tasks. Perform the following tasks in a python notebook named using the format: *Assignment-1-RollNo-Name*. Upload the python notebook to your Github Repo and submit the link to the Repo (not file) using this [Submission Form](#).

1. Compute tokens, types, and TTR (*Just present/print the stats related to each. DO NOT have a statement that prints all the tokens in the pynb.*)
2. Check if Zipf’s Law holds true for meanings and length. (*You can take some 10-20 random words (for which the meanings exists), get their meanings and work with those.*)
3. Plot Heap’s Law.