

Some key observations:

1. Most of the sentences have more than one word
2. Database Size: 4332
3. Total Number of Tokens: 19046
4. Average number of tokens in a sentence: 4.396