# Q3: **Analyze a collection of documents:**

- We know that tf-idf score enables us to find importance of words from a document in collection or corpus.

- As we can see from Q3 output, stop words like "the", "a", "an" have been excluded for most of the questions while extracting top 5 words based on their tf-idf score.

- This helps a lot in understanding important words from the document, which in turn helps in finding similar questions or differentiate dissimilar ones.