Lesson 23 (Bag of Words) Term (word) Frequency (**TF**) is the word count of a word in a document, normalized by the total number of words in the document. Consider the corpus consisting of documents that are the sentences listed below. Fill in the tables below after removing stop words. Sort words in alphabetical order. Leave zero values blank.

Document	
0	The car crashed long ago.
1	The car has rusted.
2	A rusted car is unsafe.
3	Spare car parts are needed urgently.

unique words sort in alphabetic order

don't need to fill 0's for empty

(a) Document Term Matrix:

terms (words)	ago	car	crashed	long	needed	parts	rusted	spare	unsafe	urgently
Document 0	1	1	1	1						
1		1					1			
2		1					1		1	
3		1			1	1		1		1

(b) TF (Term Frequency):

	terms (words)	ago	car	crashed	long	needed	parts	rusted	spare	unsafe	urgently
4	Document 0	1/4	1/4	1/4	1/4						
2	1		1/2					1/2			
3	2		1/3					1/3		1/3	
5	3		1/5			1/5	1/5		1/5		1/5