Exercise 2

Given is a subset of a term-document matrix which contains the term counts.
The IDF for the terms is also given. The following query is entered by a user:
„beta dome"
The following documents are relevant for the user: Dok. 2, Dok. 3.

terms/docs	ant	beta	camp	dome	end	frank	golf
Doc 1	0	3	1	0	1	1	1
Doc 2	1	2	1	3	2	2	2
Doc 3	0	0	0	1	9	1	4
Doc 4	3	4	3	0	6	0	0
Doc 5	7	8	0	0	6	4	9

	idf	log (idf)	normalized (log(idf))
ant	240	2.4	0.4
beta	120	1.2	0.2
camp	1000	3.0	0.5
dome	260000	5.4	0.9
end	900000	6.0	1
frank	4000	3.6	0.6
golf	10	1.0	1/6

a)
Calculate the ranking based on tf. (without logarithm)
Calculate a second ranking based on tf * normalized (log (idf))

b) Which ranking is better?