

A)

termo	f_{i1}	f_{i2}	f_{i3}	f_{i4}	f_{i5}	tf_{i1}	tf_{i2}	tf_{i3}	tf_{i4}	tf_{i5}	n_i	IDF	d_1	d_2	d_3	d_4	d_5
torre	1	1	0	1	0	1	1	-	1	-	3	0,736	0,736	0,736	-	0,736	-
logos	1	1	0	1	0	1	1	-	1	-	3	0,736	0,736	0,736	-	0,736	-
torre	1	0	1	0	0	1	-	1	-	-	2	1,321	1,321	-	1,321	-	-
stark	0	1	1	0	0	-	1	1	-	-	2	1,321	-	1,321	1,321	-	-
stark	0	1	1	0	0	-	1	1	-	-	2	1,321	-	1,321	1,321	-	-
Parker	0	2	0	0	0	-	2	-	-	-	1	2,585	-	4,642	-	-	-
Parker	0	2	0	0	0	-	2	-	-	-	1	2,585	-	4,642	-	-	-
Groot	0	0	1	3	2	-	-	1	2,585	2	3	0,736	-	-	0,736	1,90	1,472
Groot	0	0	1	3	2	-	-	1	-	2	2	1,321	-	-	1,321	-	2,642
Rocket	0	0	1	0	2	-	-	1	-	2	2	1,321	1,321	-	-	1,321	-
Rocket	0	0	1	0	2	-	-	1	-	2	2	1,321	1,321	-	-	1,321	-
X-MEN	1	0	0	1	0	1	-	-	1	-	2	1,321	-	1,321	2,642	-	-
X-MEN	1	0	0	1	0	1	-	-	1	-	2	1,321	-	1,321	2,642	-	-
Vingadores	0	1	2	0	0	-	1	2	-	-	2	1,321	-	-	2,642	-	1,321
Vingadores	0	1	2	0	0	-	1	2	-	-	2	1,321	-	-	2,642	-	1,321
guardiões	0	0	2	0	1	-	-	2	-	1	2	1,321	-	-	2,642	-	1,321

B) Os termos que possuem o maior TF-IDF são aqueles que são mais frequentes em cada documento e mais raros entre todos eles, como os termos Parker e Groot
(4,642) (4,108)

C) Os termos torre, stark e X-men possuem os menores valores TF-IDF, pois aparecem apenas uma vez em dois documentos. Ou seja, não são tão frequentes em cada documento e nem raros para a coleção.

D) Caso um termo apareça em todos os documentos, o IDF será igual a 0, pois $\log\left(\frac{N}{n}\right)$ será 0. Logo, o TF-IDF será 0. Isso ocorre pois, como o termo aparece em todos os documentos, sua ponderação se torna irrelevante.