Calculating Jaccard coefficient

Jaccard coefficients represent the similarity and diversity of two sets. They are calculated as the ratio between the intersection of A and B over the union of A and B. They are applied only to categorical variables, and to those for which the different levels are not considered symmetrical (e.g. Gender)

The exercise requires us to calculate Jaccard coefficients for the table below.

Name	Gender	Fever	Cough	Test-1	Test-2	Test-3	Test-4
Jack	M	Y	N	P	N	N	A
Mary	F	Y	N	P	A	P	N
Jim	M	Y	P	N	N	N	A

We are not provided a dictionary for the meanings of the different factors included, but we can assume M = Male, F = Female, Y = Yes, N = No, P = Positive/Present, and N = Negative. Y and N can therefore be recoded as 1, and N as 0. The correct meaning and appropriate coding for A is not obvious, but the hint provided indicates it should be coded as 0. In that case, we'll recode the table with Y & P = 1, and N & A = 0. Gender will not be considered as it is symmetrical.

Name	Gender	Fever	Cough	Test-1	Test-2	Test-3	Test-4
Jack	/	1	0	1	0	0	0
Mary	/	1	0	1	0	1	0
Jim	/	1	1	0	0	0	0

We are asked to calculate Jaccard coefficients for the following pairs, with results as follows:

- (Jack, Mary): (Test-3) / (Fever + Test-1 + Test-3) = 1/3 = 0.33
- (Jack, Jim): (Cough + Test-1) / (Fever + Cough + Test-1) = 2/3 = 0.67
- (Jim, Mary): (Cough + Test-1 + Test-3)/(Fever + Cough + Test-1 + Test-3) = 3/4 = 0.75