1.a

UNLess stated otherwise, submission is done individually. We rely on trust. You may

{‘Unle’, ‘nles’, ‘less’, ‘ess ‘, ‘ss s’, ‘s st’, ‘ sta’, ‘stat’, ‘tate’, ‘ated’}

1.b The statement is true. By the definition of jaccard similarity we can see that $sim(A, B) = |A \cap B| / |A \cup B|$ $hence$ $|A \cap B|$ = 0 .Now if we divide A and B to shingels according to sim(A,B) = 0 we will not be able to find any common one. Furthermore if we try any permutation of the shingels we will not have 1 at the same row for any of the shingels, which means we wont have the same minhash signature. Now by the definition of minhash we get that for two different signatures we get a value of zero. So i have shown that $sim(A, B)$ = 0 = minhash(A,B)

1.c Give an example of a two-column matrix where averaging over all the cyclic permutations does not give the Jaccard similarity. Compute Jaccard and average similarity:

Given the next two vectors (0,1,0), (0,1,1) we can get the next matrix

0 0
1 1
0 1

In this example the sets represented by the two columns have Jaccard similarity of 0.5

if we start the permutations with 1 or 2, since we meet 1 first at row 2 and then we get the same minhash. Now let's check what happens if we take 3, we get that the first 1 is at different rows which means we get different minhash value which leads to Jaccard similarity of 2/3.

2.a No teleports , and b = 0.7 by the given graph we get d=(3,2,2)

Since we have no teleports we get that A = M*b +(1-b/N) :

M=A=

1/3	1/2	0
1/3	0	1/2
1/3	1/2	1/2

Now we start with r(0) = (1/3,1/3,1/3)

We will use the power iteration method taught in class, with r(n) = Ar(n-1)

Which converges to r = (0.23,0.307,0.46)

2.b regular teleports, b = 0.85

We have same d=(3,2,2)

A = bM +(1-b/N) =

1/3	0.475	0.05
1/3	0.05	0./475

1/3	0.475	0.475

Again we start by (1/3,1/3,1/3), using the power iteration method with r(n) = Ar(n-1)
Those we get r = (0.25,0.30,0.43)

2.c b = 0.85 and teleport set is {c}
Now since we have only {c} we get A=bM +(1+b)C =

0.283	0.28	0.433
0.425	0	0.575
0	0.475	0.575

Again we start by r(0) = (1/3,1/3,1/3), r(n) = Ar(n-1)
We get r = (0.17,0.28,0.55)