Date ____ 20__

IDs assigment

Q no 1

S1 :: sunshine state enjoy sunshine
S2 :: brown fox , jump high , brown fox run
S3 : sunshine state fox run fast

Bow

	sunshine	state	enjoy	brown	fox	jump	high	run	fast
S1	2	1	1	0	0	0	0	0	0
S2	0	0	0	2	2	1	1	1	0
S3	1	1	0	0	1	0	0	1	1

TF...

	sunshine	state	enjoy	brown	fox	jump	high	run	fa
S1	2/4	1/4	1/4	0	0	0	0	0	0
S2	0	0	0	2/7	2/7	1/7	1/7	1/7	0
S3	1/5	1/5	0	0	1/5	0	0	1/5	1/

now IDF

sunshine = log(3/2) = 0.176
State = log(3/2) = 0.176
enjoy = log(3/1) = 0.477
brown = log(3/1) = 0.477

fox = log(3/2) = 0.176
jump = log(3/1) = 0.477
high = log(3/1) = 0.477
run = log(3/2) = 0.176
fast = log(3/1) = 0.477

TF-IDF

	sunshine	stated	enjoy	brown	fox	Jump	Light	run	fast
S1	0.088	0044	0.114	0	0	0	0	0	0
S2	0	0	0	0.136	0.05	0.008	0008	0.025	0
S3	0035	0035	0	0	0035	0	0	0035	0.095

vector

S1 = [2, 1, 1, 0, 0, 0, 0, 0, 0]

S2 = [0, 0, 0, 2, 2, 1, 1, 1, 0]

S3 = [1, 1, 0, 0, 1, 0, 0, 1, 1]

S1.S3 = {2+1+0+0+0+0+0+0+0} = 3

$|S1| = (2^2 + 1^2 + 1^2 + 0+0+0+0+0+0+0)^{1/2}$

$|S1| = (6)^{1/2} = 2.45$

$|S3| = (1^2 + 1^2 + 0 + 0 + 1^2 + 0 + 0 + 1^2 + 1^2)^{1/2} = 2.24$

$\frac{3}{(2.45)(2.24)} = \cos \quad = 0.55$