Homework 1

Question 1:

Input Matrix

	a	b	c	d	e	f	g	h
A	4	5	0	5	1	0	3	2
B	0	3	4	3	1	2	1	0
C	2	0	1	3	0	4	5	3

Treat the utility matrix as Boolean:

Jaccard Distance		
	A	B
B	0.5	
C	0.5	0.5

Cosine Distance		
	A	B
B	0.33	
C	0.33	0.33

Use a different discretization:

Jaccard Distance		
	A	B
B	0.6	
C	0.67	0.83

Cosine Distance		
	A	B
B	0.42	
C	0.5	0.71

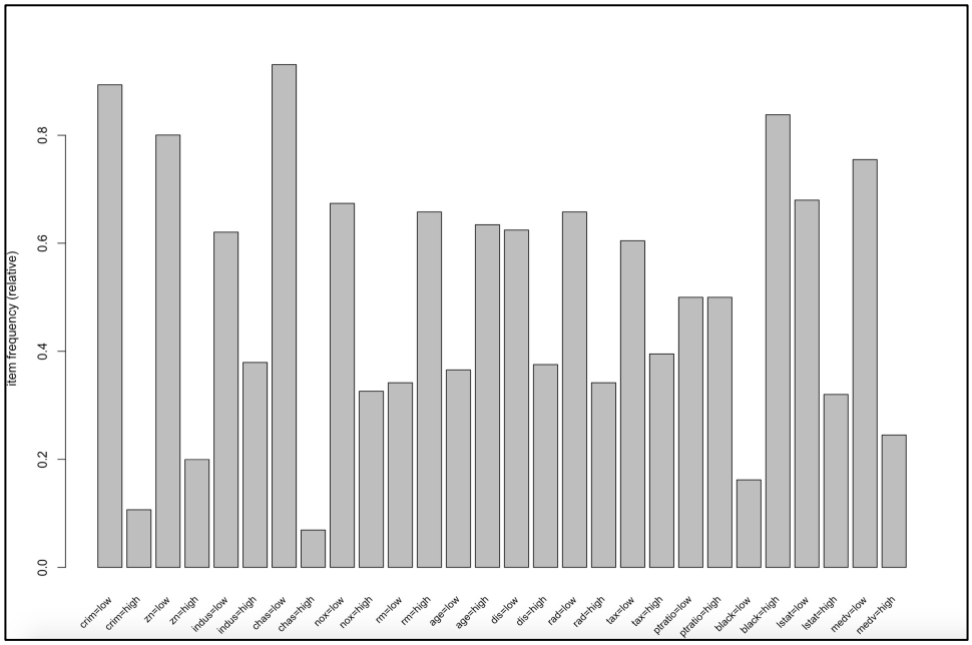
Normalize the matrix by subtracting from each nonblank entry the average value for its user:

Jaccard Distance		
	A	B
B	0.5	
C	0.5	0.5

Cosine Distance		
	A	B
B	0.36	
C	0.34	0.33

Question 2:

Item Frequency Plot



Student is interested is a low crime area as close to the city as possible the association rules suggests the area having following features:

Feature	**Value**
Nitrogen oxide level	High
Near Charles River	High (Yes)
Average number of rooms per dwelling	Low
Proportion of non-retail business acres per town	High
Property Tax rate	High
Proportion of blacks by town	Low

For a family wanting area with school with low pupil-teacher ratio the association rules suggests the area having following features:

Feature	**Value**
Nitrogen oxide level	Low
Near Charles River	High (Yes)
Average number of rooms per dwelling	High
Proportion of non-retail business acres per town	Low
Property Tax rate	Low
Proportion of blacks by town	Low
Accessibility to radial highways	Low