

Age	Sex	Zipcode	Disease
5	M	12000	gastric ulcer
9	M	14000	dyspepsia
6	M	18000	pneumonia
8	M	19000	bronchitis
12	M	22000	pneumonia
15	M	25000	pneumonia
19	M	24000	pneumonia
21	F	58000	flu
26	F	36000	gastritis
28	F	37000	flu
56	F	33000	flu

Name	Age	Sex	Zipcode
Andy	5	M	12000
Bill	9	M	14000
Ken	6	M	18000
Nash	8	M	19000
Joe	12	M	22000
George	15	M	25000
Sam	19	M	24000
Linda	21	F	58000
Jame	26	F	36000
Sarah	28	F	37000
Mary	56	F	33000

Generalization

Age : [1,10] [11-20] [21-60]

Zipcode : [10001-15000] [15001-20000] [20001-30000] [30001-60000]

Age	Sex	Zipcode	Disease	lloss
[1-10]	M	[10001-15000]	gastric ulcer	0.25
[1-10]	M	[10001-15000]	dyspepsia	0.25
[1-10]	M	[15001-20000]	pneumonia	0.25
[1-10]	M	[15001-20000]	bronchitis	0.25
[11-20]	M	[20001-30000]	pneumonia	0.35
[11-20]	M	[20001-30000]	pneumonia	0.35
[11-20]	M	[20001-30000]	pneumonia	0.35
[21-60]	F	[30001-60000]	flu	1.25
[21-60]	F	[30001-60000]	gastritis	1.25
[21-60]	F	[30001-60000]	flu	1.25
[21-60]	F	[30001-60000]	flu	1.25

7.05

K-anonymity			
QIDs		k-values	min k value
[1-10]	M	[10001-15000]	2
[1-10]	M	[15001-20000]	2
[11-20]	M	[20001-30000]	3
[21-60]	F	[30001-60000]	4

L-diversity		
QIDs	L-values	min L value

[1-10]	M	[10001-15000]	2	1
[1-10]	M	[15001-20000]	2	
[11-20]	M	[20001-30000]	1	
[21-60]	F	[30001-60000]	2	

Entropy L-diversity					
QIDs			QIDs logs	L values	min L value
[1-10]	M	[10001-15000]	0.3010	2.0	1
[1-10]	M	[15001-20000]	0.3010	2.0	
[11-20]	M	[20001-30000]	0	1.0	
[21-60]	F	[30001-60000]	0.2442	1.8	

Age : [1-20] [20-60]

Zipcode : [10001-20000] [20001-30000] [30001-60000]

Age	Sex	Zipcode	Disease	lloss
[1-20]	M	[10001-20000]	gastric ulcer	0.52
[1-20]	M	[10001-20000]	dyspepsia	0.52
[1-20]	M	[10001-20000]	pneumonia	0.52
[1-20]	M	[10001-20000]	bronchitis	0.52
[1-20]	M	[20001-30000]	pneumonia	0.52
[1-20]	M	[20001-30000]	pneumonia	0.52
[1-20]	M	[20001-30000]	pneumonia	0.52
[21-60]	F	[30001-60000]	flu	1.25
[21-60]	F	[30001-60000]	gastritis	1.25
[21-60]	F	[30001-60000]	flu	1.25
[21-60]	F	[30001-60000]	flu	1.25

8.62

K-anonymity				
QIDs			k-values	min k value
[1-20]	M	[10001-20000]	4	3
[1-20]	M	[20001-30000]	3	
[21-60]	F	[30001-60000]	4	

L-diversity				
QIDs			L-values	min L value
[1-20]	M	[10001-20000]	4	1
[1-20]	M	[20001-30000]	1	
[21-60]	F	[30001-60000]	2	

Entropy L-diversity					
QIDs			QIDs logs	L values	min L value
[1-20]	M	[10001-20000]	0.6021	4.0	1
[1-20]	M	[20001-30000]	0	1.0	
[21-60]	F	[30001-60000]	0.2442	1.8	

Anatomization

Age	Sex	Zipcode	Group ID
5	M	12000	1
9	M	14000	1
6	M	18000	1
8	M	19000	1
12	M	22000	2
15	M	25000	2

19	M	24000	2
21	F	58000	3
26	F	36000	3
28	F	37000	3
56	F	33000	3

Group ID	Disease	Count
1	gastric ulcer	1
1	dyspepsia	1
1	pneumonia	1
1	bronchitis	1
2	pneumonia	3
3	flu	3
3	gastritis	1

How many people aged 26 are expected to have flu?

Generalized	$1/40 * 3$	0.075
Anatomized	$1/4 * 3$	0.75
Real value		1

Ο πίνακας Α πρόκειται να δημοσιοποιηθεί από ένα οργανισμό υγείας
Ο πίνακας Β είναι δημοσίως διαθέσιμος

Age	Sex	Zipcode	Disease
5	M	12000	gastric ulcer
9	M	14000	dyspepsia
6	M	18000	pneumonia
8	M	19000	bronchitis
12	M	22000	pneumonia
15	M	25000	pneumonia
19	M	24000	pneumonia
21	F	58000	flu
26	F	36000	gastritis
28	F	37000	flu
56	F	33000	flu

Name	Age	Sex	Zipcode
Andy	5	M	12000
Bill	9	M	14000
Ken	6	M	18000
Nash	8	M	19000
Joe	12	M	22000
George	15	M	25000
Sam	19	M	24000
Linda	21	F	58000
Jame	26	F	36000
Sarah	28	F	37000
Mary	56	F	33000

Υπολογίστε τα:

- k για το k-anonymity
- I για I-diversity
- I για entropy I-diversity

Για τα παρακάτω generalizations υπολογίστε :

- k για το k-anonymity
- I για I-diversity
- I για entropy I-diversity
- και Iloss

Age : [1,10] [11-20] [21-60]

Zipcode : [10001-15000] [15001-20000] [20001-30000] [30001-60000]

Age : [1-20] [20-60]

Zipcode : [10001-20000] [20001-30000] [30001-60000]

Age	Sex	Zipcode	Disease	lloss
[1-10]	M	[10001-15000]	gastric ulcer	0.25
[1-10]	M	[10001-15000]	dyspepsia	0.25
[1-10]	M	[15001-20000]	pneumonia	0.25
[1-10]	M	[15001-20000]	bronchitis	0.25
[11-20]	M	[20001-30000]	pneumonia	0.35
[11-20]	M	[20001-30000]	pneumonia	0.35
[11-20]	M	[20001-30000]	pneumonia	0.35
[21-60]	F	[30001-60000]	flu	1.25
[21-60]	F	[30001-60000]	gastritis	1.25
[21-60]	F	[30001-60000]	flu	1.25
[21-60]	F	[30001-60000]	flu	1.25

7.05

K-anonymity				
QIDs			k-values	min k value
[1-10]	M	[10001-15000]	2	2
[1-10]	M	[15001-20000]	2	
[11-20]	M	[20001-30000]	3	
[21-60]	F	[30001-60000]	4	

L-diversity				
QIDs			L-values	min L value
[1-10]	M	[10001-15000]	2	1
[1-10]	M	[15001-20000]	2	
[11-20]	M	[20001-30000]	1	
[21-60]	F	[30001-60000]	2	

Entropy L-diversity					
QIDs			QIDs logs	L values	min L value
[1-10]	M	[10001-15000]	0.3010	2.0	1
[1-10]	M	[15001-20000]	0.3010	2.0	
[11-20]	M	[20001-30000]	0	1.0	
[21-60]	F	[30001-60000]	0.2442	1.8	