

## Excercises “Database Privacy”

Below table provides health information of patients for medical studies.

Date of Birth	Zip code	Allergy	History of Illness
24-03-78	71767	penicillin	pharyngitis
02-07-69	70728	no allergy	stroke
11-03-78	71767	no allergy	polio
07-03-59	70730	penicillin	diabetes mellitus
02-07-69	70730	sulfur	diphtheria
01-07-60	71767	no allergy	colitis

### Excercise 1:

You assume that two of your neighbors are in above table:

1. Bob – he is 31 years old and lives in a town with zip code 71767;
2. Alice – she is 40 years old and lives in a town with zip code 70728.

What can you learn about their diseases?

### Excercise 2:

Make above table 2-anonymous by using

1. generalization only,
2. supression only,
3. a mixture of both.

Which Data Generalization Hierarchy DGH did you use?

### Excercise 3:

Did you create a 2-anonymous table with minimal loss of information?

### Excercise 4:

Can you construct a table which includes Alice’s record and satisfies 2-anonymity but where you are still able to infer the illness of Alice?