

Privacy-Preserving Data Publishing

Exercises : Partition-Based Models and Algorithms

1 Privacy Models : k -Anonymity and l -Diversity

The well-known parisien health center *Brascassés* wants to share his patients' health data with epidemiologists. Before starting sharing data, the DBA of Brascassés would like to ask to an expert his opinion on the privacy guarantees of the sanitization methods used. You are the expert.

The dataset (Figure 1) is a relation with three columns and about 20 tuples. Each tuple corresponds to a different patient. The DBA has set the QI to (**Age**, **Code postal**) and the SD to **Diagnostic**. Figures 2 to 5 represent possible sanitized releases of the dataset, denoted G_i below. All are candidates to being published.

1. For each release, what are the equivalence classes? Give a name to each class (e.g., EC_i).
2. For the top left class of G_3 , what is the set of sensitive data?
3. Draw on a 2D graph the equivalence classes of G_3 . You can use the graphs given in the appendix, where QIs are already printed.
4. Which release satisfy k -Anonymity with : $k = 3$? $k = 5$? $k = 8$?
5. In general, is an 8-Anonymous release also 5-Anonymous?
6. Does G_3 satisfy naive l -Diversity with $l = 3$.
7. Bob is a patient of Brascassés. His tuple appears in the initial dataset and is equal to (75003, 37, Oblomovisme). Give Bob's possible diagnostics for an attacker receiving G_3 and knowing :
 - (a) Nothing;
 - (b) That Bob participates to the initial dataset
 - (c) *Idem 7b* and that his QI is (75003, 37);
 - (d) *Idem 7c* and that Bob does not have « Grippe »;
 - (e) *Idem 7c* and that Alice and Charlie (1) have participated to the dataset, (2) have QIs respectively equal to (75002, 35) and (75004, 34), and (3) have SDs respectively equal to Grippe and Rhinopharyngite;
8. Assuming that releases G_3 and G_4 are both published. Which privacy criteria does each one satisfy independently from the other? What are the guarantees after observing the two together? (You can focus on the top left classes only.)
9. (*Option, Hard*) Observing several releases of the same dataset occurs when several updates are published (insertions, deletions, updates). Imagine a privacy criteria that supports the publishing of a dataset that evolves over time.

Dataset, Releases, and Graphs

	Code postal	Age	Diagnostic
1	75010	22	Angine rouge
2	75001	36	Grippe
3	75007	35	Angine blanche
4	75003	24	Grippe
5	75008	24	Rhinopharyngite
6	75002	35	Grippe
7	75001	32	Rhinopharyngite
8	75003	37	Oblomovisme
9	75001	38	Angine
10	75004	23	Grippe
11	75005	25	Grippe
12	75008	34	Angine blanche
13	75003	23	Grippe
14	75006	24	Rhinopharyngite
15	75009	30	Angine blanche
16	75008	32	Angine
17	75006	31	Angine érythémateuse
18	75004	28	Rhinopharyngite
19	75004	34	Rhinopharyngite
20	75001	20	Grippe
21	75007	38	Angine érythémato-pultacée

FIGURE 1 – Dataset

Code postal *	Age *	Diagnostic
[75005, 75010]	[20, 40]	Angine rouge
[75001, 75004]	[20, 40]	Grippe
[75005, 75010]	[20, 40]	Angine blanche
[75001, 75004]	[20, 40]	Grippe
[75005, 75010]	[20, 40]	Rhinopharyngite
[75001, 75004]	[20, 40]	Oblomovisme
[75001, 75004]	[20, 40]	Rhinopharyngite
[75001, 75004]	[20, 40]	Grippe
[75001, 75004]	[20, 40]	Angine
[75001, 75004]	[20, 40]	Grippe
[75005, 75010]	[20, 40]	Grippe
[75005, 75010]	[20, 40]	Angine blanche
[75001, 75004]	[20, 40]	Grippe
[75005, 75010]	[20, 40]	Rhinopharyngite
[75005, 75010]	[20, 40]	Angine blanche
[75005, 75010]	[20, 40]	Angine
[75005, 75010]	[20, 40]	Angine érythémateuse
[75001, 75004]	[20, 40]	Rhinopharyngite
[75001, 75004]	[20, 40]	Rhinopharyngite
[75001, 75004]	[20, 40]	Grippe
[75005, 75010]	[20, 40]	Angine érythémato-pultacée

FIGURE 2 – Release G₁

Code postal *	Age *	Diagnostic
[75005, 75010]	[20, 30]	Angine rouge
[75001, 75004]	[31, 40]	Grippe
[75005, 75010]	[31, 40]	Angine blanche
[75001, 75004]	[20, 30]	Grippe
[75005, 75010]	[20, 30]	Rhinopharyngite
[75001, 75004]	[31, 40]	Oblomovisme
[75001, 75004]	[31, 40]	Rhinopharyngite
[75001, 75004]	[31, 40]	Grippe
[75001, 75004]	[31, 40]	Angine
[75001, 75004]	[20, 30]	Grippe
[75005, 75010]	[20, 30]	Grippe
[75005, 75010]	[31, 40]	Angine blanche
[75001, 75004]	[20, 30]	Grippe
[75005, 75010]	[20, 30]	Rhinopharyngite
[75005, 75010]	[20, 30]	Angine blanche
[75005, 75010]	[31, 40]	Angine
[75005, 75010]	[31, 40]	Angine érythémateuse
[75001, 75004]	[20, 30]	Rhinopharyngite
[75001, 75004]	[31, 40]	Rhinopharyngite
[75001, 75004]	[20, 30]	Grippe
[75005, 75010]	[31, 40]	Angine érythémato-pultacée

FIGURE 3 – Release G₂

Code postal *	Age *	Diagnostic
[75008, 75010]	[20, 30]	Angine rouge
[75001, 75001]	[31, 40]	Grippe
[75005, 75007]	[31, 40]	Angine blanche
[75001, 75003]	[20, 30]	Grippe
[75008, 75010]	[20, 30]	Rhinopharyngite
[75002, 75004]	[31, 40]	Oblomovisme
[75001, 75001]	[31, 40]	Rhinopharyngite
[75002, 75004]	[31, 40]	Grippe
[75001, 75001]	[31, 40]	Angine
[75004, 75004]	[20, 30]	Grippe
[75005, 75007]	[20, 30]	Grippe
[75008, 75010]	[31, 40]	Angine blanche
[75001, 75003]	[20, 30]	Grippe
[75005, 75007]	[20, 30]	Rhinopharyngite
[75008, 75010]	[20, 30]	Angine blanche
[75008, 75010]	[31, 40]	Angine
[75005, 75007]	[31, 40]	Angine érythémateuse
[75004, 75004]	[20, 30]	Rhinopharyngite
[75002, 75004]	[31, 40]	Rhinopharyngite
[75001, 75003]	[20, 30]	Grippe
[75005, 75007]	[31, 40]	Angine érythémato-pultacée

FIGURE 4 – Release G₃

Code postal *	Age *	Diagnostic
[75008, 75010]	[20, 30]	Angine rouge
[75001, 75004]	[31, 36]	Grippe
[75005, 75007]	[31, 40]	Angine blanche
[75001, 75003]	[20, 30]	Grippe
[75008, 75010]	[20, 30]	Rhinopharyngite
[75001, 75004]	[31, 36]	Grippe
[75001, 75004]	[31, 36]	Rhinopharyngite
[75001, 75004]	[37, 40]	Oblomovisme
[75001, 75004]	[37, 40]	Angine
[75004, 75004]	[20, 30]	Grippe
[75005, 75007]	[20, 30]	Grippe
[75008, 75010]	[31, 40]	Angine blanche
[75001, 75003]	[20, 30]	Grippe
[75005, 75007]	[20, 30]	Rhinopharyngite
[75008, 75010]	[20, 30]	Angine blanche
[75008, 75010]	[31, 40]	Angine
[75005, 75007]	[31, 40]	Angine érythémateuse
[75004, 75004]	[20, 30]	Rhinopharyngite
[75001, 75004]	[31, 36]	Rhinopharyngite
[75001, 75003]	[20, 30]	Grippe
[75005, 75007]	[31, 40]	Angine érythémato-pultacée

FIGURE 5 – Release G₄









