

Homework 2

Name: Julian Steiner

Matriculation No.: 2669944

Problem 1

(1.1)

Table 1: 3-anonymit Crime Data by State in Germany

(id)	Name	Age	Gender	State of Germany	Crime
1	*	30-39	Male	Rhineland-Palatinate	Murder
2	*	20-29	Female	Bavaria	Murder
3	*	30-39	Female	North Rhine-Westphalia	Robbery
4	*	30-39	Male	Rhineland-Palatinate	Assault
5	*	20-29	Female	Bavaria	Robbery
6	*	30-39	Female	North Rhine-Westphalia	Murder
7	*	30-39	Male	Rhineland-Palatinate	Parking
8	*	10-19	Male	Hesse	Murder
9	*	30-39	Female	North Rhine-Westphalia	Parking
10	*	20-29	Female	Bavaria	Speeding
11	*	10-19	Male	Hesse	Robbery
12	*	30-39	Male	North Rhine-Westphalia	Assault
13	*	30-39	Male	North Rhine-Westphalia	Speeding
14	*	10-19	Male	Hesse	Speeding
15	*	30-39	Male	North Rhine-Westphalia	Murder

(1.2)

- Crime: Murder, Assault, Parking \Rightarrow l-diversity \Rightarrow 3

Table 2: Group 1 - Crime Data by State in Germany

(id)	Name	Age	Gender	State of Germany	Crime
1	*	30-39	Male	Rhineland-Palatinate	Murder
4	*	30-39	Male	Rhineland-Palatinate	Assault
7	*	30-39	Male	Rhineland-Palatinate	Parking

- Crime: Murder, Robbery, Speeding \Rightarrow l-diversity \Rightarrow 3

Table 3: Group 2 - Crime Data by State in Germany

(id)	Name	Age	Gender	State of Germany	Crime
2	*	20-29	Female	Bavaria	Murder
5	*	20-29	Female	Bavaria	Robbery
10	*	20-29	Female	Bavaria	Speeding

- Crime: Robbery, Murder, Parking \Rightarrow l-diversity \Rightarrow 3

Table 4: Group 3 - Crime Data by State in Germany

(id)	Name	Age	Gender	State of Germany	Crime
3	*	30-39	Female	North Rhine-Westphalia	Robbery
6	*	30-39	Female	North Rhine-Westphalia	Murder
9	*	30-39	Female	North Rhine-Westphalia	Parking

- Crime: Murder, Robbery, Speeding \Rightarrow l-diversity \Rightarrow 3

Table 5: Group 4 - Crime Data by State in Germany

(id)	Name	Age	Gender	State of Germany	Crime
8	*	10-19	Male	Hesse	Murder
11	*	10-19	Male	Hesse	Robbery
14	*	10-19	Male	Hesse	Speeding

- Crime: Assault, Speeding, Murder \Rightarrow l-diversity \Rightarrow 3

Table 6: Group 5 - Crime Data by State in Germany

(id)	Name	Age	Gender	State of Germany	Crime
12	*	30-39	Male	North Rhine-Westphalia	Assault
13	*	30-39	Male	North Rhine-Westphalia	Speeding
15	*	30-39	Male	North Rhine-Westphalia	Murder

Note: From the above equivalence classes, each class has 3 unique crimes. Therefore, the l-diversity of the modified table is 3.

(1.3)

Table 7: Group 1 - Crime Data by State in Germany

(id)	Name	Age	Gender	State of Germany	Crime
1	*	30-39	Male	Rhineland-Palatinate	Murder
4	*	30-39	Male	Rhineland-Palatinate	Assault
7	*	30-39	Male	Rhineland-Palatinate	Parking

- **Q → General Distribution**

- Assault → $2/15 \rightarrow 0.133$
- Murder → $5/15 \rightarrow 0.333$
- Parking → $2/15 \rightarrow 0.133$
- Robbery → $3/15 \rightarrow 0.2$
- Speeding → $3/15 \rightarrow 0.2$

$$Q = (0.133, 0.333, 0.133, 0.2, 0.2)$$

- **P → Distribution of equivalence class**

- Assault → $1/3 \rightarrow 0.333$
- Murder → $1/3 \rightarrow 0.333$
- Parking → $1/3 \rightarrow 0.333$
- Robbery → $0/3 \rightarrow 0.0$
- Speeding → $0/3 \rightarrow 0.0$

$$P = (0.333, 0.333, 0.333, 0.0, 0.0)$$

- **t-closeness**

$$\begin{aligned}
 D(P, Q) &= \sum_{i=1}^m \frac{1}{2} |p_i - q_i| = \\
 &= \frac{1}{2} * (|0.133 - 0.333| + |0.333 - 0.333| + |0.133 - 0.333| + |0.2 - 0.0| + |0.2 - 0.0|) = \\
 &= \frac{1}{2} * 0.8 = 0.4
 \end{aligned}$$

Problem 2

See jupyter notebook for function implementation. The results of the function shown in 8.

p	X_0	X_1	Y_0	Y_1
0.0	0.8	0.2	0.20	0.80
0.2	0.8	0.2	0.32	0.68
0.5	0.8	0.2	0.50	0.50
0.8	0.8	0.2	0.68	0.32
1.0	0.8	0.2	0.80	0.20

Table 8

- **Privacy**

- **p=0 and p=1:**
 - * When $p = 0$, the noisy distribution Y is exactly the reverse of the true distribution X . This provides no privacy because the distribution Y directly reveals the opposite of X .
 - * When $p = 1$, the noisy distribution Y is identical to the true distribution X . This also provides no privacy because Y is exactly X .
- **p=0.5:** When $p = 0.5$, the noisy distribution Y becomes uniform (i.e., $Y = [0.5, 0.5]$). This provides the highest level of privacy because Y does not reveal any information about the true distribution X .
- **p=0.2 and p=0.8:** These intermediate values provide a balance between privacy and utility. As p moves away from 0.5 towards 0 or 1, the privacy decreases because Y starts resembling X more closely.

- **Utility**

- **p=0 and p=1:** These values provide maximum utility because Y is either exactly X or exactly the reverse of X . There is a clear correspondence between Y and X , making Y highly useful for accurate analysis.
- **p=0.5:** When $p = 0.5$, the utility is the lowest because the noisy distribution Y is uniform and does not reflect the true distribution X at all. This makes Y less useful for analysis.
- **p=0.2 and p=0.8:** These intermediate values provide a balance between privacy and utility. As p moves away from 0.5 towards 0 or 1, the utility increases because Y starts resembling X more closely.

Problem 3

• Text 1

Had such a blast hanging out with [NAME] and [NAME] in #[LOC] [DATE]! We explored the city, went to the #[LOC]Zoo, and had an awesome time. It's always a good time catching up with old friends and making new memories together. [LOC] has so much character – definitely need to come back soon! Thanks for the awesome day, [NAME] and [NAME]. Let's do it again sometime! #Friends #[LOC] #FB20 #TUDA

• Text 2

Date: [DATE]

To: [URL], [URL], [URL], [URL]

Title: RE: RE: RE: website

Dear website team,

here is a kick-off mail for the website. Format – a github page which lists some papers, between [URL] (too little info) and [URL] (too much for now). For starters, we'd pick 50 papers from the list we already made [URL]. [NAME] has kindly agreed to lead the work on this. [NAME] has agreed to develop the site. [NAME] and [NAME] have agreed to help with selecting papers and giving feedback. We need to have it running by [DATE] to link it in the paper. With that, I pass it to [NAME].

I guess the first step is setting up the repo and adding all people in this thread as Maintainers. My Github account is [URL]. If there are any questions, please let me know, if something urgent pops up, give me a call at [PHONE].

Best,