Friday 25th January, 2019

1 SQL Stuff

Definition 1. *Let* $\mathcal{W} = (\mathcal{C} \cup \times \mathcal{F} \cup (\times \mathcal{F})^2, \approx, \leqslant, \Pi, \sigma, \ltimes, 1, 2, 3, ...)$ *where*

- $\mathcal{C}$ *is a set of domain names*
- $\mathcal{F}$ *is a set of domain sets*
- $\approx$ *is a relation that compares entries*
- $\leqslant$ *is a relation that compares cardinalities*
- $\Pi, \sigma, \ltimes$ *are usual relational algebraic functions*
- $1, 2, 3, ...$ *are constant sets with cardinalities* $1, 2, 3, ...$

A database $\mathcal{D} \in (\times \mathcal{F})^2$ is an embedding in $\mathcal{W}$.

Definition 2 (k-Anonymity). *For* $k \in \{1, 2, 3, ...\}$ *and* $\mathcal{Q} \subseteq \mathcal{C}$*, a database* $\mathcal{D}$ *is* k*-anonymous in regards to* $\mathcal{Q}$ *iff it holds that*

$$\forall_{d \in \mathcal{D}} \exists_{S \subseteq \mathcal{D}} \quad (S \geqslant k \wedge \forall_{s \in S} \quad \Pi_{\mathcal{Q}} s \approx \Pi_{\mathcal{Q}} d)$$

A SQL query to discover all the values that violate k-Anonymity is

```
SELECT Q, COUNT(*) AS count FROM D GROUP BY Q HAVING count < k
```

Definition 3 (ℓ-Diversity). *For* $l \in \{1, 2, 3, ...\}$ *and* $\mathcal{R} \subseteq \mathcal{C} - \mathcal{Q}$*, a database* $\mathcal{D}$ *that's* k*-anonymous in regards to* $\mathcal{Q}$ *is* ℓ*-diverse in regards to* $\mathcal{R}$ *iff it holds that*

$$\forall_{d \in \mathcal{D}} \exists_{S \subseteq \mathcal{D}} \quad (S \geqslant k \wedge \forall_{s \in S} \quad \Pi_{\mathcal{Q}} s \approx \Pi_{\mathcal{Q}} d \wedge \Pi_{\mathcal{R}} S \geqslant \ell)$$

A SQL query to discover all the values that violate ℓ-Diversity is

```
SELECT Q, COUNT(*) AS count FROM
    (SELECT Q FROM D GROUP BY Q,R) AS T
GROUP BY Q HAVING count < ℓ
```

Definition 4 ((X,Y)-Privacy). *For $k \in \{1,2,3,...\}$ and $X, Y \subseteq \mathcal{C}$, a database is (X,Y)-anonymous for k iff it holds that*

$$\forall_{d \in \mathcal{D}} \exists_{S \subseteq \mathcal{D}} \ (\Pi_Y S \geqslant k \wedge \forall_{s \in S} \ \Pi_X s \approx \Pi_X d)$$

```
SELECT X, count(*) AS count FROM
    (SELECT X, count(*) FROM D GROUP BY X,Y)
AS T GROUP BY X HAVING count < k
```