

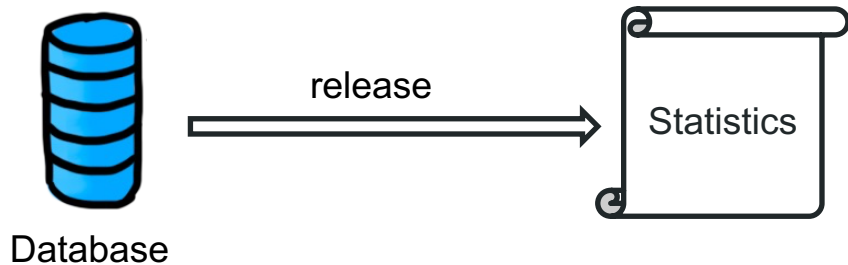
CS459/698

Privacy, Cryptography, Network and Data Security

Differential Privacy

Spring 2025, Monday/Wednesday 2:30pm-4:50pm

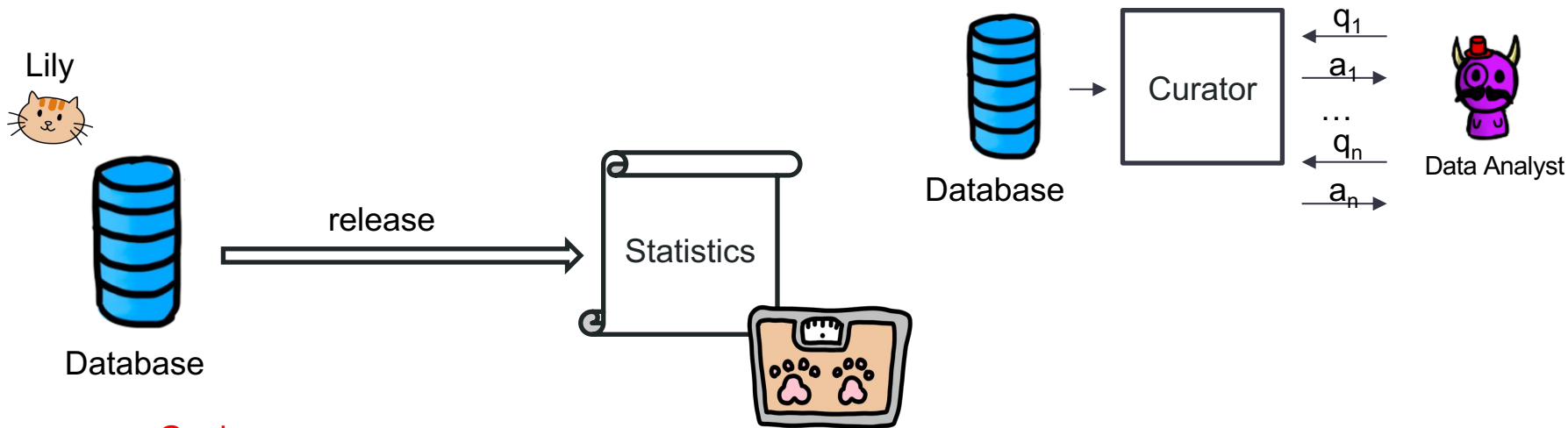
The Background: Privacy-Preserving Data Analysis



Goal:

- Protect privacy
- Provide useful information (utility)

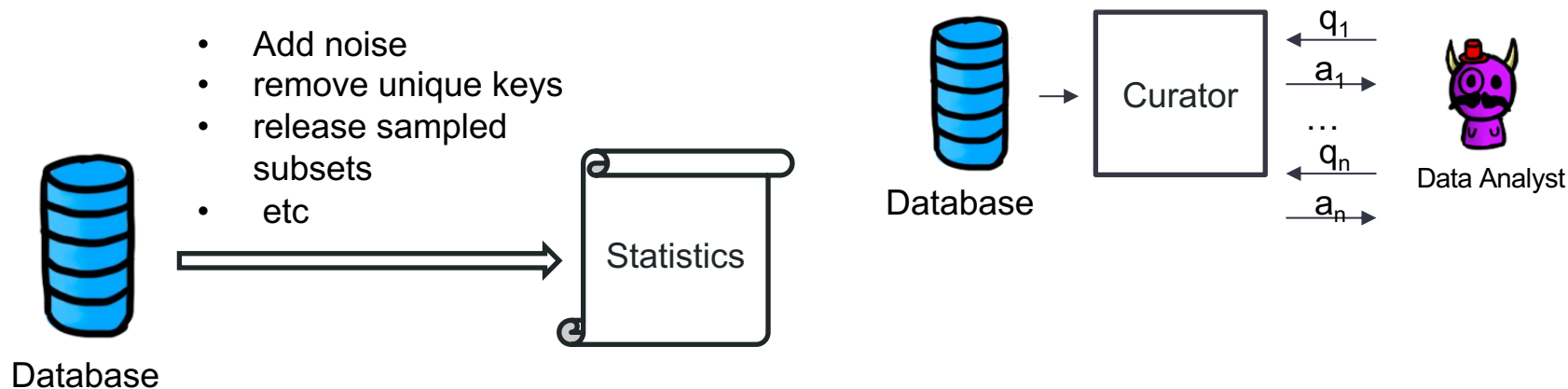
The Background: Privacy-Preserving Data Analysis



Goal:

- Protect privacy
- Provide useful information (utility)

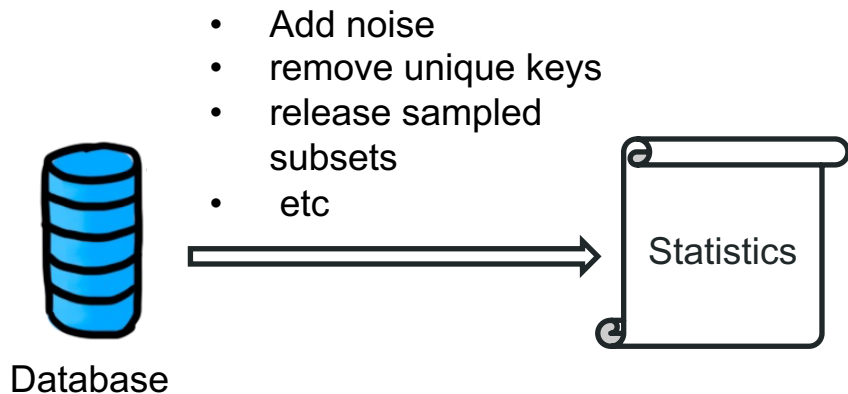
The Background: Privacy-Preserving Data Analysis



Goal:

- Protect privacy
- Provide useful information (utility)

The Background: Privacy-Preserving Data Analysis

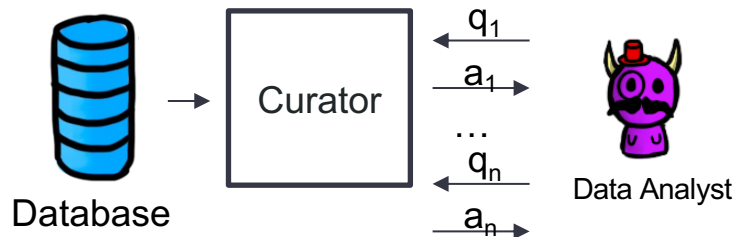


Goal:

- Protect privacy
- Provide useful information (utility)

- Anything that can be learned about an individual from a statistical database can be learned without access to the database. [Dalenius 1977]

- Adversary: querier



What is privacy?

What is Privacy? Privacy, considered as part of our system of values, is not an isolated freedom: if privacy changes, much else of the whole structure of human interaction and values will change [1]. While privacy is regarded as a human right, providing a clear definition of privacy is challenging. Different cultures emphasize different aspects of privacy, and legal concepts of privacy have varied over time. In 1890, Warren and Brandeis defined the right to privacy as what Judge Cooley called “the right to be let alone” [2]. This definition was an extension of the concept of “the right to life” [2, 3]. In 1960, Prosser concluded that the law of privacy comprises four distinct kinds of invasion of four different interests of the plaintiff: intrusion, public disclosure of private facts, false light in the public eye, and appropriation [4]. In 1992, Bennett pointed out the threat posed to privacy by the ease of personal information collection during the information revolution [5]. Over a decade later, Westin further defined privacy as “the claim of an individual to determine what information about himself or herself should be known to others” [6]. While the traditional method of conceptualizing privacy focuses on locating its essence to form a single generic label, in Solove’s 2002 work, a privacy invasion was first defined as the interference with the integrity of certain practices, and privacy was then defined as a general term to describe the practices to be protected [7].

We define privacy as the right to protect oneself against four categories of violations, based on the taxonomy of privacy proposed by Solove [8, 9]:

Information Collection: The right to protect one’s information from being gathered.

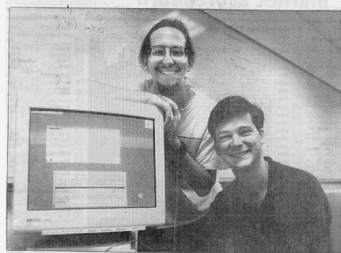
Information Processing: The right to protect one’s information from being stored, analyzed, and manipulated.

Information Dissemination: The right to decide whether and how one’s information is transferred.

Invasion: The right to protect oneself from intrusion or direct interference in decision-making.

[1] A. Warren, “Privacy and the Right to Be Let Alone,” *Harvard Law Review*, vol. 4, no. 5, pp. 509–520, Dec. 1890.
[2] S. D. Warren and L. D. Brandeis, “The Right to Privacy,” *Harvard Law Review*, vol. 4, no. 5, pp. 509–520, Dec. 1890.
[3] L. Ashworth, “What is Privacy? The History and Definition of Privacy,” 2002.
[4] W. L. Prosser, “Privacy,” *California Law Review*, vol. 48, no. 3, pp. 389–403, Aug. 1960.
[5] C. J. Bennett, “Regulating Privacy: Data Protection and Public Policy in Europe and the United States,” 1992.
[6] A. F. Westin, “Toward a Political and Social Theory of Privacy,” *Journal of Social Issues*, vol. 48, no. 2, pp. 419–430, Jul. 1992.
[7] D. J. Solove, “Conceptualizing Privacy,” *California Law Review*, vol. 90, no. 4, pp. 1067–1098, Jul. 2002.
[8] —, “A Taxonomy of Privacy,” *University of Pennsylvania Law Review*, vol. 154, no. 3, pp. 475–501, Jan. 2006.
[9] —, “Five Gut Holdings to Help and Other Misunderstandings of Privacy,” *San Diego Law Review*, vol. 44, pp. 745–772, Jul. 2007.

Cyberpunks: activists advocating strong cryptography since late 1980s



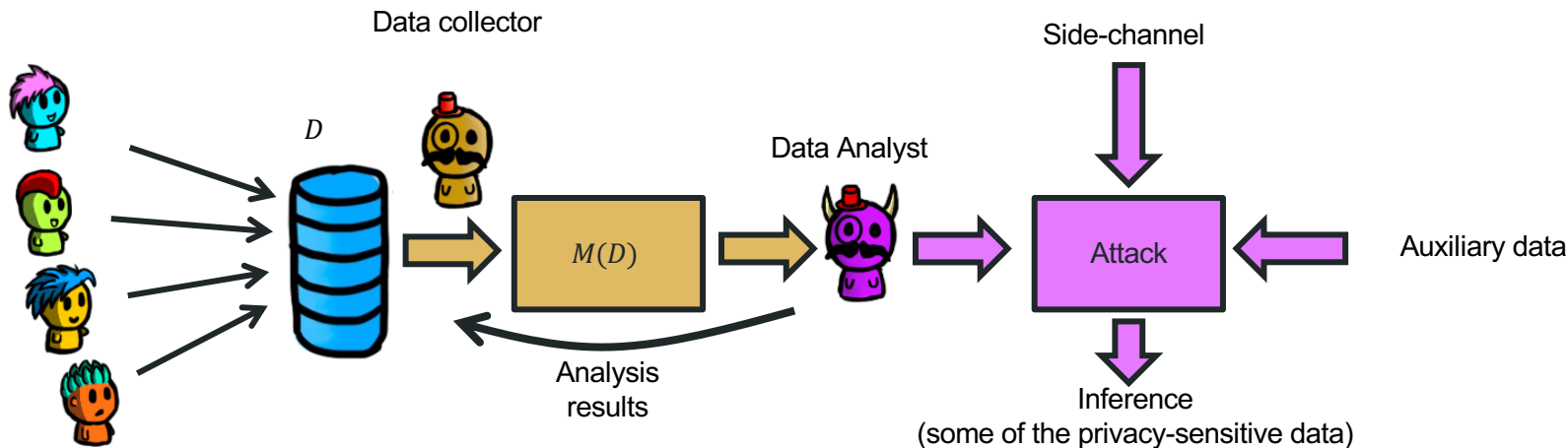
University of California Cyberpunks Ian Goldberg and David Wagner cracked Netscape’s security code.

The Cyberpunks are not only worried about the government spying on electronic messages. With the explosion in goods and services being sold the Internet, the group focused its efforts on illuminating the security risks of the electronic world. On Sept. 17, 1995, Ian Goldberg and David Wagner, both graduate students at the University of California at Berkeley, said they had cracked Netscape’s security code. Image source: Boston Sunday Globe, 8 October, 1995, pp. B1, B7. Accessed via <https://people.eecs.berkeley.edu/~daw/press/ian/ian4.html>

- Cryptography aims for strongest possible privacy: **indistinguishability**
- **Semantic security [Goldwasser–Micali 1982]**
 - Anything that can be learned from an encrypted message can be learned without access to the encrypted message
 - Adversary: an eavesdropper

Can we protect against auxiliary information?

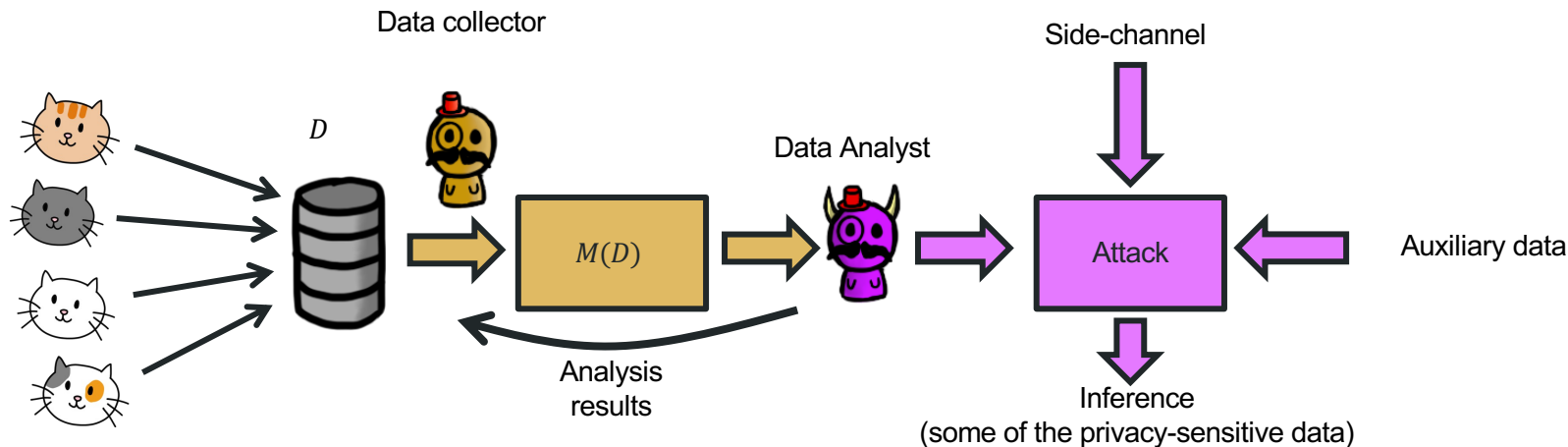
- Each user contributes to one entry (row) of a database D .
- The release mechanism M publishes some data $R = M(D)$.
 - Note: we can characterize the mechanism by $\Pr(M(D) = R)$, which is the same as $\Pr(R|D)$ on inference attacks
- Can we provide privacy when the adversary has **auxiliary information**?



Can we protect against auxiliary information?

- Lily is 0.5 kg heavier than the average weight of a cat in Canada.
- If we know the average weight of a cat in Canada, we can know Lily's weight

Q: Can we design a mechanism M that prevents this? Does it make sense to design a mechanism M that prevents this?



Can we protect against auxiliary information?

- Lily is 0.5 kg heavier than the average weight of a cat in Canada.
- If we know the average weight of a cat in Canada, we can know Lily's weight

Q: Can we design a mechanism M that prevents this? Does it make sense to design a mechanism M that prevents this?

A: The adversary would reach the same conclusion even if Lily isn't in the database! We cannot prevent this unless we destroy utility.

Can we protect against auxiliary information?

- Lily is 0.5 kg heavier than the average weight of a cat in Canada.
- If we know the average weight of a cat in Canada, we can know Lily's weight

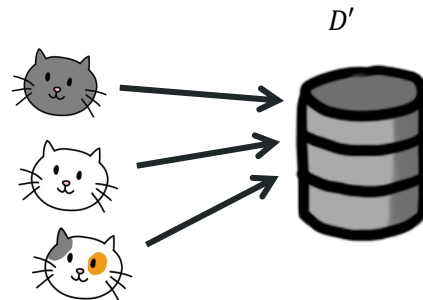
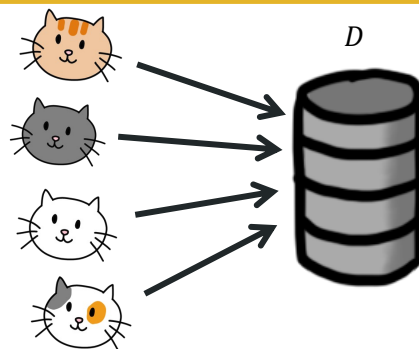
Q: Can we design a mechanism M that prevents this? Does it make sense to design a mechanism M that prevents this?

A: The adversary would reach the same conclusion even if Lily isn't in the database! We cannot prevent this unless we destroy utility.

The only way to achieve the perfect notions of privacy is no statistics at all

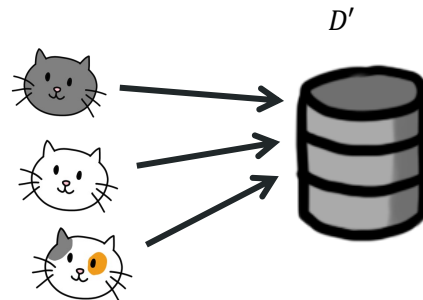
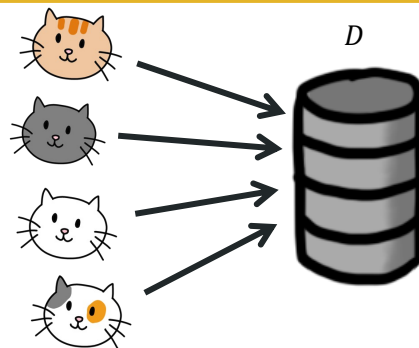
Differential Privacy: A Definition

- We need a new definition!
- It captures the contribution of Lily to the statistics.



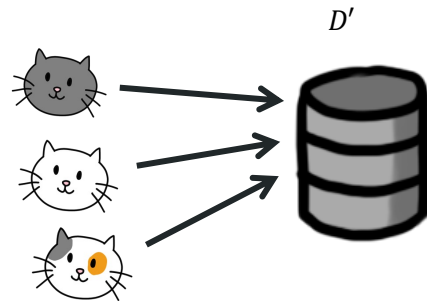
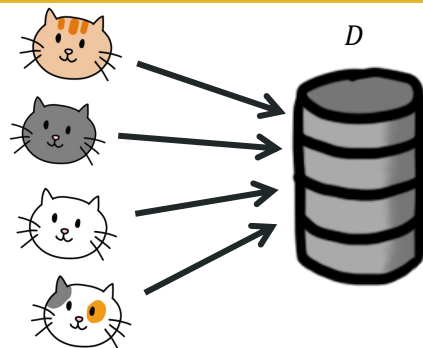
Differential Privacy: A Definition

- Differential privacy is a definition of privacy tailored to the problem of privacy-preserving data analysis
 - If the analyst learns similar things in these two cases about Lily, then M provides enough privacy
 - If the adversary learns “a lot” about Lily in both cases, then we cannot prevent this anyway



Differential Privacy: A Definition

- Differential privacy is a definition of privacy tailored to the problem of privacy-preserving data analysis
 - Given $R = M(D)$, the adversary should be unable to distinguish whether or not Lily was in the dataset!
 - Note that this means that $M(D)$ has to be randomized (or always report the same value, but this makes R constant – independent of D – which is not useful.)
- One should set an appropriate amount of noise depending on each particular use case.
 - We want to preserve data privacy. We don't want to destroy utility



The data collectors' argument

... on trying to persuade you to join a differentially private survey:

- *You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available. (bla... differential privacy ... bla)*

The data collectors' argument

... on trying to persuade you to join a differentially private survey:

- *You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available. (bla... differential privacy ... bla)*
- But this is only true if they tell you what algorithm they use to release your data and you have verified that their algorithm is indeed differentially private.

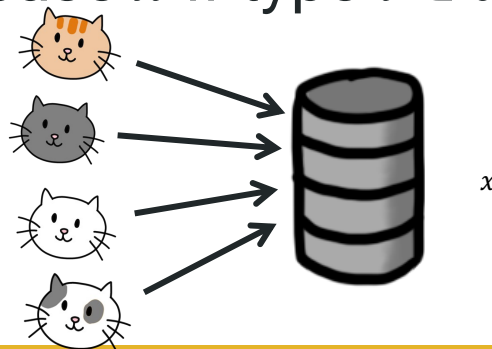
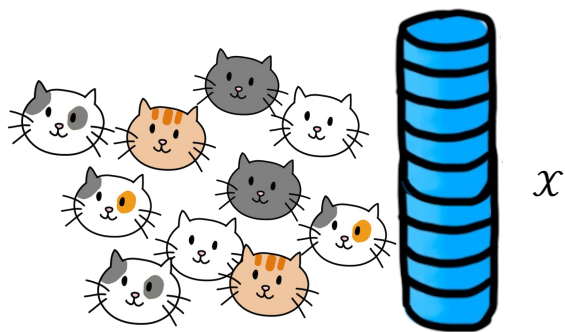
The data collectors' argument

... on trying to persuade you to join a differentially private survey:

- *You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available. (bla... differential privacy ... bla)*
- But this is only true if they tell you what algorithm they use to release your data and you have verified that their algorithm is indeed differentially private.
- Parameters/details matter a lot... as we will see




Universe

- Think of database x as being collections of records from a universe \mathcal{X}
- It is convenient to represent databases by their histograms $x \in \mathbb{N}^{|\mathcal{X}|}$, in which each entry x_i represents the number of elements in the database x if type $i \in \mathcal{X}$.



Universe

- Think of database x as being collections of records from a universe \mathcal{X}
- It is convenient to represent databases by their histograms $x \in \mathbb{N}^{|\mathcal{X}|}$, in which each entry x_i represents the number of elements in the database x if type $i \in \mathcal{X}$.

Lily	Canada	Orange	
Gracie	U.K.	Blue	
Luna	Canada	Tuxedo	
...	

\mathcal{X}

Lily	Canada	Orange	
Luna	Canada	Tuxedo	
...	

$$x = \{1, 0, 1 \dots\}$$



ℓ_1 distance

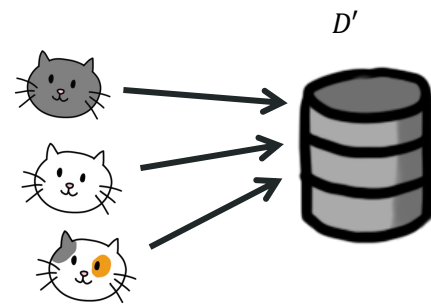
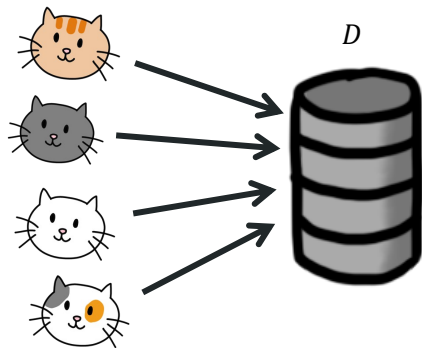
The ℓ_1 norm of a database x is denoted $\|x\|_1$ and is defined to be: $\|x\|_1 = \sum_{i=1}^{|X|} |x_i|$

The ℓ_1 distance between two databases x and y is $\|x - y\|_1$

- Here, $x = \{x_1, x_2, \dots\}$, $y = \{y_1, y_2, \dots\}$
- $\|x - y\|_1$ is the measure of how many records differ between x and y

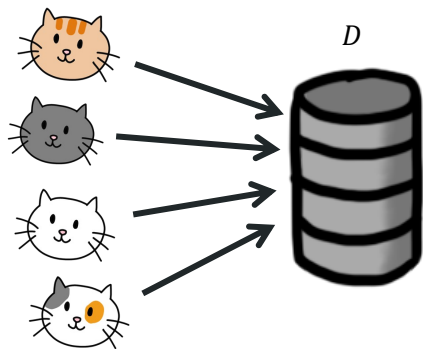
Neighboring datasets

- Assume for now that the databases differ on one single record
- These datasets are usually called **neighboring datasets** (and usually denoted by D and D')

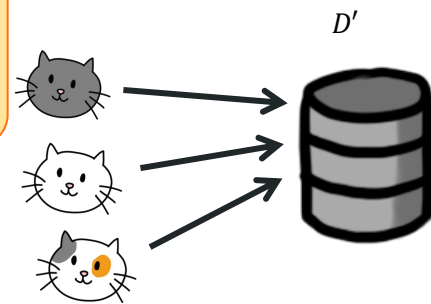


Neighboring datasets

- Assume for now that the databases differ on one single record
- These datasets are usually called **neighboring datasets** (and usually denoted by D and D')

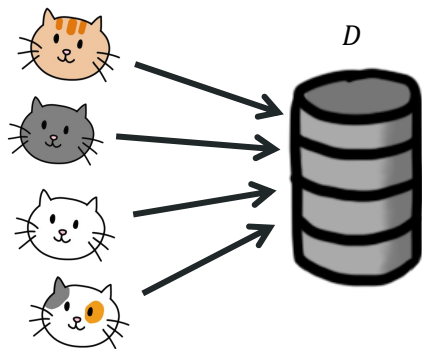


Q: What's the ℓ_1 distance between the neighboring datasets D and D' (with and without Lily)?



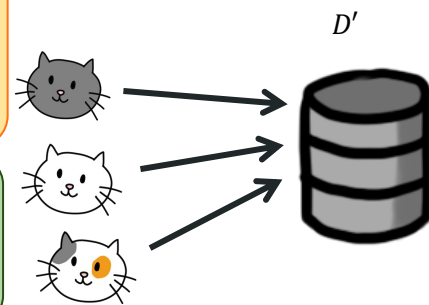
Neighboring datasets

- Assume for now that the databases differ on one single record
- These datasets are usually called **neighboring datasets** (and usually denoted by D and D')



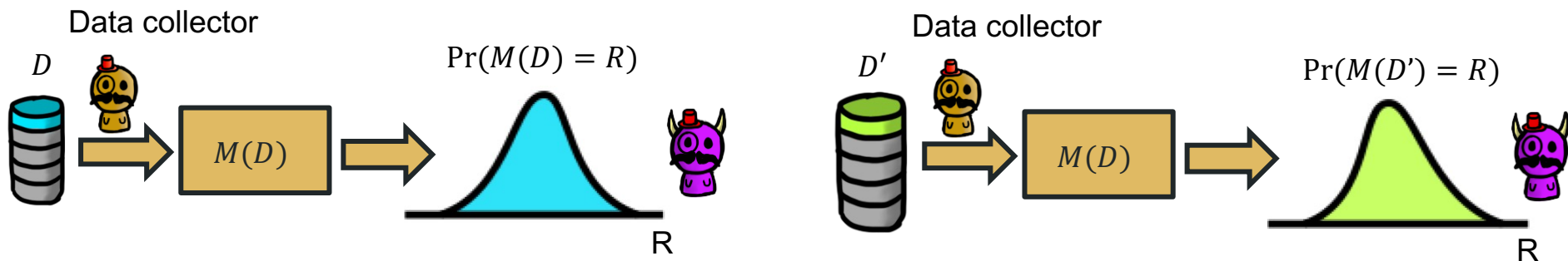
Q: What's the ℓ_1 distance between the neighboring datasets D and D' (with and without Lily)?

A: $\|D - D'\|_1 = 1$



Back on topic: We want similar output distributions!

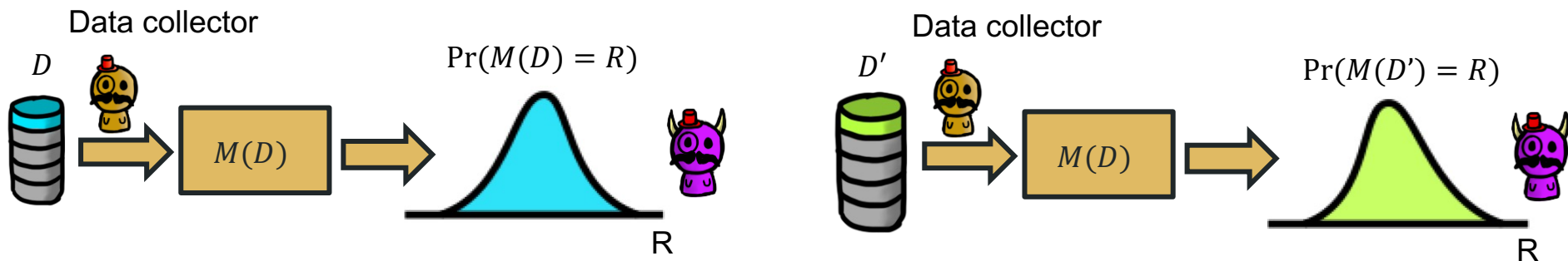
- We want these distributions to be “similar” (for all R)
- If the mechanism M behaves **nearly identically** for D and D' , then an attacker can't tell whether D or D' was used (and hence can't learn much about the individual).



Back on topic: We want similar output distributions!

- We want these distributions to be “similar” (for all R)
- If the mechanism M behaves **nearly identically** for D and D' , then an attacker can't tell whether D or D' was used (and hence can't learn much about the individual).

Q: How do we quantify this similarity?



How do we define “similar” distributions?

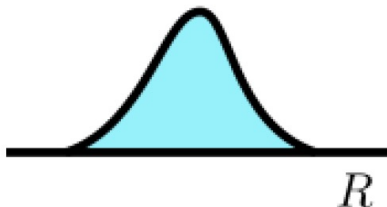
Tentative privacy definition (with parameter p)

A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

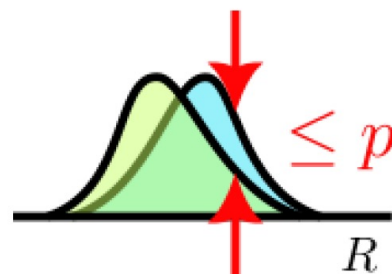
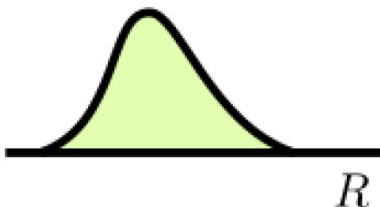
$$\Pr(M(D') = R) - p < \Pr(M(D) = R) < \Pr(M(D') = R) + p$$

- What does this mean?

$\Pr(M(D) = R)$



$\Pr(M(D') = R)$



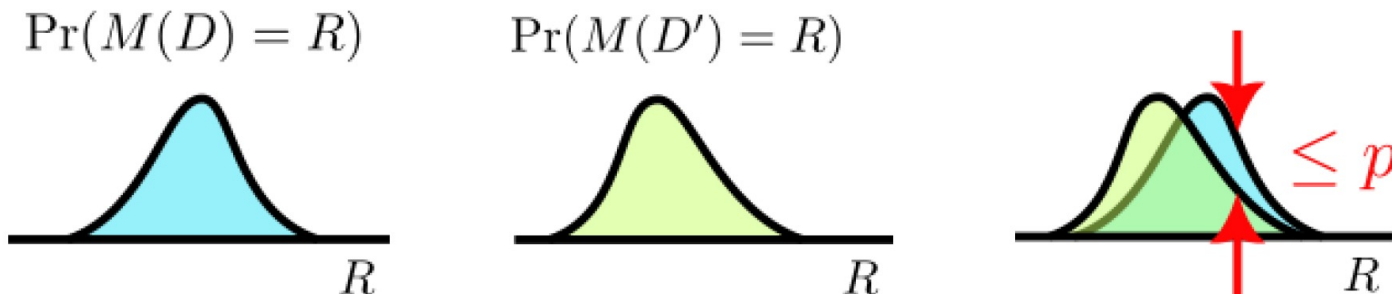
How do we define “similar” distributions?

Tentative privacy definition (with parameter p)

A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

$$\Pr(M(D') = R) - p < \Pr(M(D) = R) < \Pr(M(D') = R) + p$$

- What does this mean?



M is a randomized algorithm, so not a single output but a distribution of outputs

How do we define “similar” distributions?

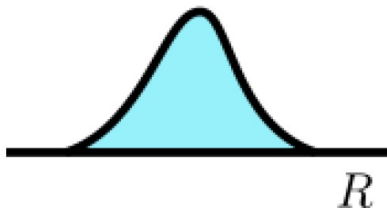
Tentative privacy definition (with parameter p)

A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

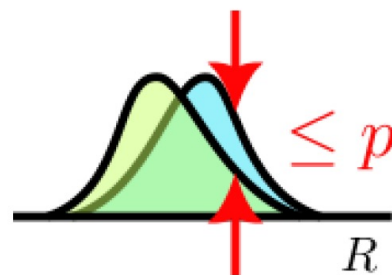
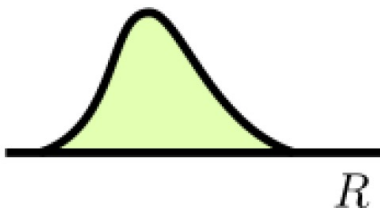
$$\Pr(M(D') = R) - p < \Pr(M(D) = R) < \Pr(M(D') = R) + p$$

- What does this mean?

$\Pr(M(D) = R)$



$\Pr(M(D') = R)$



Q: What gives more privacy, small or large p ?

How do we define “similar” distributions?

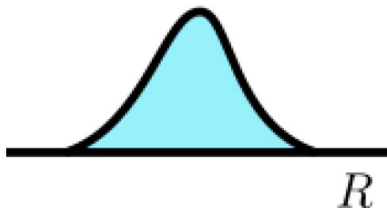
Tentative privacy definition (with parameter p)

A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

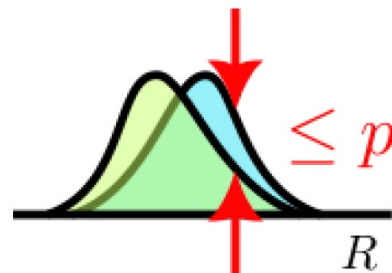
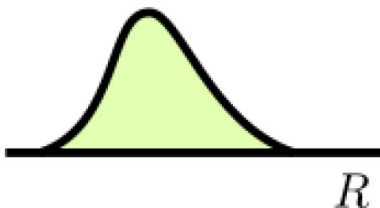
$$\Pr(M(D') = R) - p < \Pr(M(D) = R) < \Pr(M(D') = R) + p$$

- What does this mean?

$\Pr(M(D) = R)$



$\Pr(M(D') = R)$



Q: What gives more privacy, small or large p ?

A: Small p , the distributions are more alike

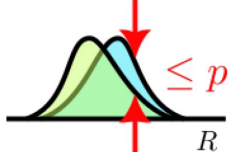
Does this really work?

Tentative privacy definition (with parameter p)

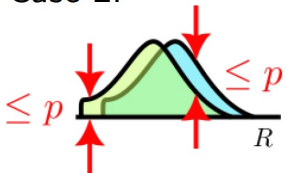
A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

$$\Pr(M(D') = R) - p < \Pr(M(D) = R) < \Pr(M(D') = R) + p$$

Case 1:



Case 2:



Q: Case 1 seems fine. What is the issue with case 2?

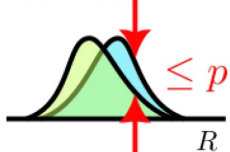
Does this really work?

Tentative privacy definition (with parameter p)

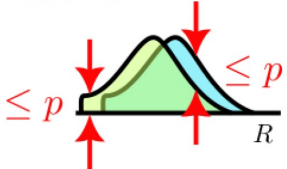
A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

$$\Pr(M(D') = R) - p < \Pr(M(D) = R) < \Pr(M(D') = R) + p$$

Case 1:



Case 2:



Q: Case 1 seems fine. What is the issue with case 2?

A: There are some outputs R that can only happen if the input was D' (e.g., if Alice was not in the dataset). This allows the adversary to distinguish between D and D' with 100% certainty. In other words, the attacker can find a **perspective** through which the two databases behave differently.

What if we make the distance multiplicative?

Tentative privacy definition II (with parameter p)

A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

$$\frac{\Pr(M(D') = R)}{p} < \Pr(M(D) = R) < \Pr(M(D') = R) \cdot p$$



Q: what does provide more privacy, small (but larger than 1) or large p ?

What if we make the distance multiplicative?

Tentative privacy definition II (with parameter p)

A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

$$\frac{\Pr(M(D') = R)}{p} < \Pr(M(D) = R) < \Pr(M(D') = R) \cdot p$$



Q: what does provide more privacy, small (but larger than 1) or large p ?

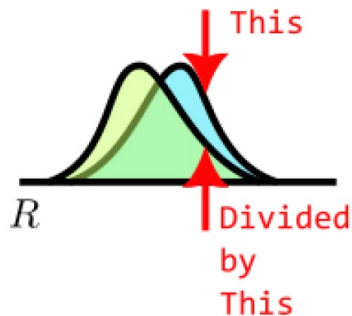
A: Small p

What if we make the distance multiplicative?

Tentative privacy definition II (with parameter p)

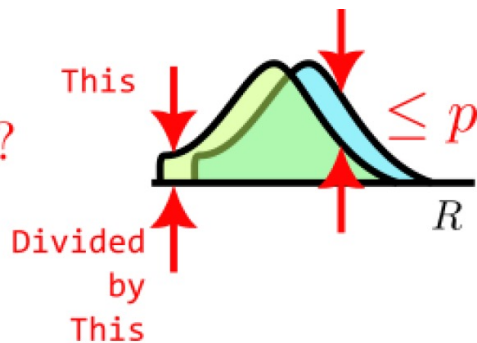
A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

$$\frac{\Pr(M(D') = R)}{p} < \Pr(M(D) = R) < \Pr(M(D') = R) \cdot p$$



Q: Does this make sense?

$\leq \infty?$

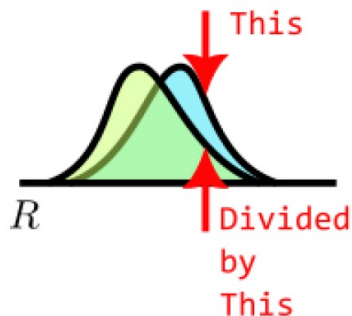


What if we make the distance multiplicative?

Tentative privacy definition II (with parameter p)

A mechanism M is p -private if the following holds for all possible outputs R and all pairs of neighboring datasets (D, D') :

$$\frac{\Pr(M(D') = R)}{p} < \Pr(M(D) = R) < \Pr(M(D') = R) \cdot p$$

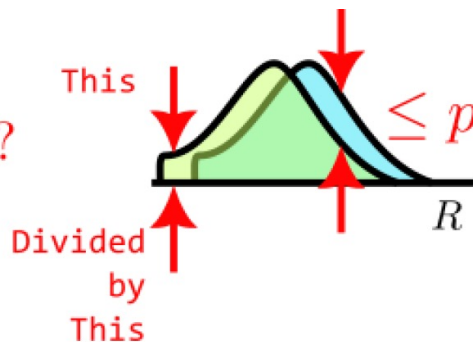


Q: Does this make sense?

$\leq p$

$\leq \infty?$

A: Yes, because this is the case where we get no privacy, and that's what $p = \infty$ means



Differential Privacy

- Same definition, but instead of “ p ” we use e^ϵ

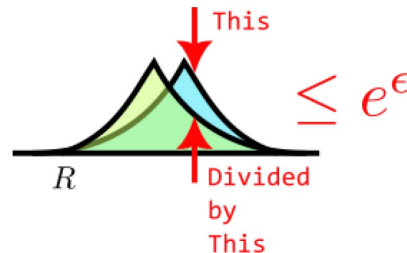
Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is ϵ -differentially private (ϵ -DP) if the following holds for all possible outputs $R \in \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

$$\Pr(M(D) = R) \leq \Pr(M(D') = R) e^\epsilon$$

Some notes:

- We use e^ϵ , instead of just ϵ , because this makes it easier to formulate some useful theorems
- We do not need the $e^{-\epsilon}$ on the left, since this must hold for all pairs (D, D') . This includes (D', D) .
- $\epsilon \in [0, \infty)$; this ensures that $e^\epsilon \in [1, \infty)$

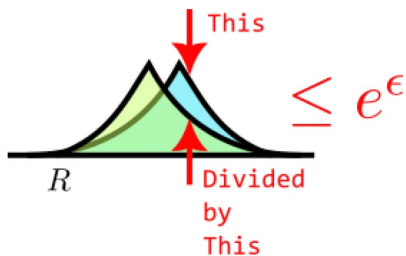


Differential Privacy

Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is ϵ -differentially private (ϵ -DP) if the following holds for all possible outputs $R \in \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

$$\Pr(M(D) = R) \leq \Pr(M(D') = R) e^\epsilon$$



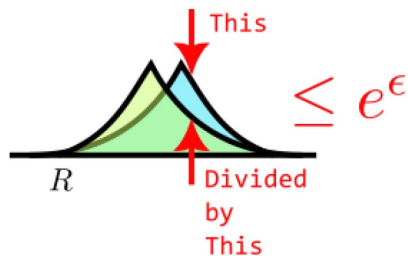
Q: which provides more privacy? $\epsilon = 1$ or $\epsilon = 2$?

Differential Privacy

Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is ϵ -differentially private (ϵ -DP) if the following holds for all possible outputs $R \in \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

$$\Pr(M(D) = R) \leq \Pr(M(D') = R) e^\epsilon$$



Q: which provides more privacy? $\epsilon = 1$ or $\epsilon = 2$?

A: Smaller ϵ means more privacy; larger means less privacy

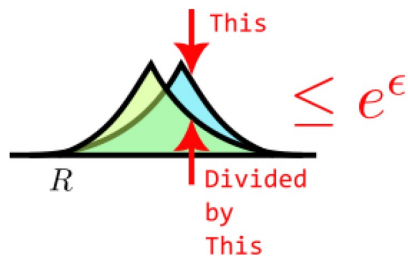
Q: What does $\epsilon = 0$ mean?

Differential Privacy

Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is ϵ -differentially private (ϵ -DP) if the following holds for all possible outputs $R \in \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

$$\Pr(M(D) = R) \leq \Pr(M(D') = R) e^{\epsilon}$$



Q: which provides more privacy? $\epsilon = 1$ or $\epsilon = 2$?

A: Smaller ϵ means more privacy; larger means less privacy

Q: What does $\epsilon = 0$ mean?

A: Perfect privacy! The output is independent of the dataset!
Utility will be very bad.

Some notes on Differential Privacy

- DP was proposed in 2006 by Cynthia Dwork et al.
[\[DMNS06\]](#)
- The authors won the Test-of-Time Award in 2016 and the Godel Price in 2017.
- Adopted by big tech like Apple, Google, Microsoft, Facebook, LinkedIn, and by the US Census Bureau for the 2020 US Census
- There is no consensus on how small ϵ should be.



Should I share my data?

Recall the data collectors' argument

- There is no consensus on how small ϵ should be.

You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available. (bla... differential privacy... bla)



No ϵ information!!!

Should I share my data?



Recall the data collectors' argument

- There is no consensus on how small ϵ should be.

You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available. (bla... differential privacy... bla). We have $\epsilon = 4$



Is the ϵ small enough?

Should I share my data?



Recall the data collectors' argument

- There is no consensus on how small ϵ should be.

You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available. (bla... differential privacy... bla)
We have $\epsilon = 4$



ϵ is unit-less & contextless
It provides probabilistic guarantees
[Nanayakkara 2023]



I am just a kitty cat.
I don't understand!

DP Mechanisms

or in other words, how to add noise and how much?

Sensitivity

- Q: How much noise to add? → Measure sensitivity!

Sensitivity

- Q: How much noise to add? → Measure sensitivity!
- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

Sensitivity

- **Q:** How much noise to add? → Measure sensitivity!
- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} \|f(D) - f(D')\|_1$$

- **Note 1:** The range of f is k -dimensional
 - e.g., Avg. and Sum. of different attributes in a public data release
- **Note 2:** ℓ_1 -sensitivity is the ℓ_1 -norm

$$\|\vec{x}_1 - \vec{x}_2\|_1 = \sum_i |\vec{x}_1[i] - \vec{x}_2[i]|$$

Sensitivity

- **Q:** How much noise to add? → Measure sensitivity!
- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} \|f(D) - f(D')\|_1$$

- **Note 1:** The range of f is k -dimensional
 - e.g., Avg. and Sum. of different attributes in a public data release
- **Note 2:** ℓ_1 -sensitivity is the ℓ_1 -norm

$$\|\vec{x}_1 - \vec{x}_2\|_1 = \sum_i |\vec{x}_1[i] - \vec{x}_2[i]|$$

It captures the magnitude by which a single individual's data can change the function f in the worst case.

Sensitivity

- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$



- How many cats in the database are orange? 🐱
- How many in the database orange 🐱, and how many are blue 🐱?
- How many are heavier than 4.5 kg? 🐱
- How many are orange, and how many are heavier than 4.5 kg? 🐱

Sensitivity

- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$



- How many cats in the database are orange? 🐱 $\Delta_1=1$
- How many in the database orange 🐱, and how many are blue 🐱?
- How many are heavier than 4.5 kg? 🐱
- How many are orange, and how many are heavier than 4.5 kg? 🐱

Sensitivity

- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$



- How many cats in the database are orange? 🐱 $\Delta_1=1$
- How many in the database orange 🐱, and how many are blue 🐱? $\Delta_1=1$
- How many are heavier than 4.5 kg? 🐱
- How many are orange, and how many are heavier than 4.5 kg? 🐱

Sensitivity

- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$



- How many cats in the database are orange? 🐱 $\Delta_1=1$
- How many in the database orange 🐱, and how many are blue 🐱? $\Delta_1=1$
- How many are heavier than 4.5 kg? 🐱 $\Delta_1=1$
- How many are orange, and how many are heavier than 4.5 kg? 🐱

Sensitivity

- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -**sensitivity** of f is the maximum change that replacing D for D' can cause in the output:

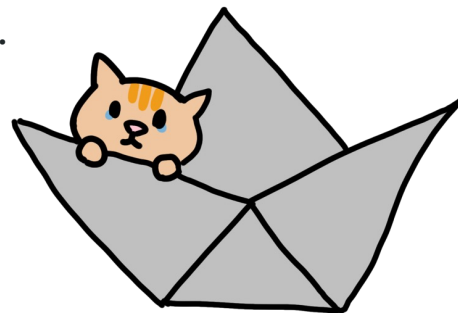
$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$



- How many cats in the database are orange? 🐱 $\Delta_1=1$
- How many in the database orange 🐱, and how many are blue 🐱? $\Delta_1=1$
- How many are heavier than 4.5 kg? 🐱 $\Delta_1=1$
- How many are orange, and how many are heavier than 4.5 kg? 🐱 $\Delta_1=2$

Sensitivity w/ one pair of neighboring databases

- There is a very evil STEM professor, asking you to build a paper boat that can take a certain weight and test your boat with cats! (He hates cats since an orange cat called Lily stole all the dried sardines at his home one day.)
- Even though, as a Cat Privacy activist to protect cats from weight surveillance, you decide to protect the poor cats' privacy.
- We assume a cat's weight is between 0 kg and 5 kg.

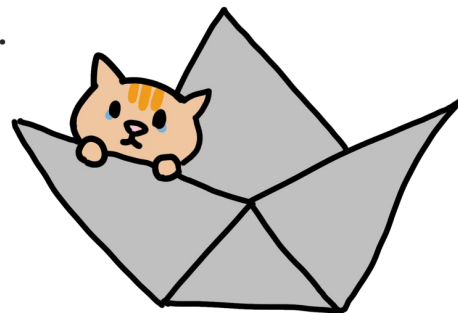


Sensitivity w/ one pair of neighboring databases

- There is a very evil STEM professor, asking you to build a paper boat that can take a certain weight and test your boat with cats! (He hates cats since an orange cat called Lily stole all the dried sardines at his home one day.)
- Even though, as a Cat Privacy activist to protect cats from weight surveillance, you decide to protect the poor cats' privacy.
- We assume a cat's weight is between 0 kg and 5 kg.

We have a dataset **D** with Lily **included**:

- Lily: 5 kg
- Every cat else: 0 ~ 5 kg



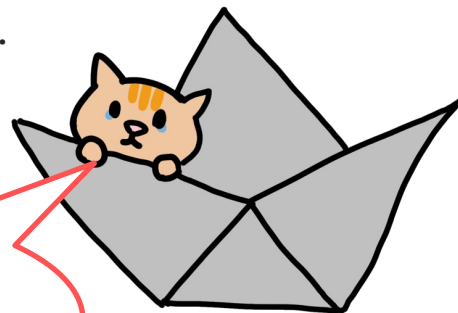
Sensitivity w/ one pair of neighboring databases

- There is a very evil STEM professor, asking you to build a paper boat that can take a certain weight and test your boat with cats! (He hates cats since an orange cat called Lily stole all the dried sardines at his home one day.)
- Even though, as a Cat Privacy activist to protect cats from weight surveillance, you decide to protect the poor cats' privacy.
- We assume a cat's weight is between 0 kg and 5 kg.

We have a dataset **D** with Lily **included**:

- Lily: 5 kg
- Every cat else: 0 ~ 5 kg

I am the
heaviest
kitty! Help!



Sensitivity w/ one pair of neighboring databases

- We assume a cat's weight is between 0 kg and 5 kg.
- **D** with Lily **included**:
 - Lily: 5 kg
 - Every cat else: 0 ~ 5 kg

Algorithm: You are allowed to make a query that returns the sum of the weights of the cats in the database.

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

It captures the magnitude by which a single individual's data can change the function f in **the worst case**.



Sensitivity w/ one pair of neighboring databases

- **D** with Lily **included**:

- Lily: 5 kg
- Every cat else: 0 ~ 5 kg

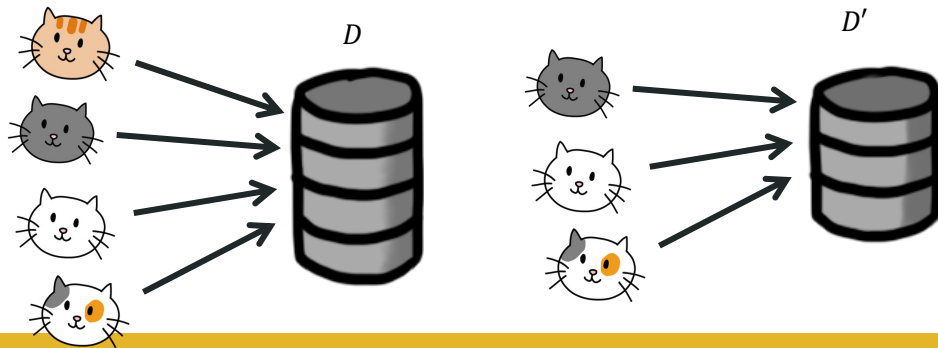
- **D'** with Lily **not included**:

- Every cat else: 0 ~ 5 kg

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

Algorithm: You are allowed to make a query that returns the sum of the weights of the cats in the database.

Q: What is the ℓ_1 -sensitivity here?



Sensitivity w/ one pair of neighboring databases

- **D** with Lily **included**:

- Lily: 5 kg
- Every cat else: 0 ~ 5 kg

- **D'** with Lily **not included**:

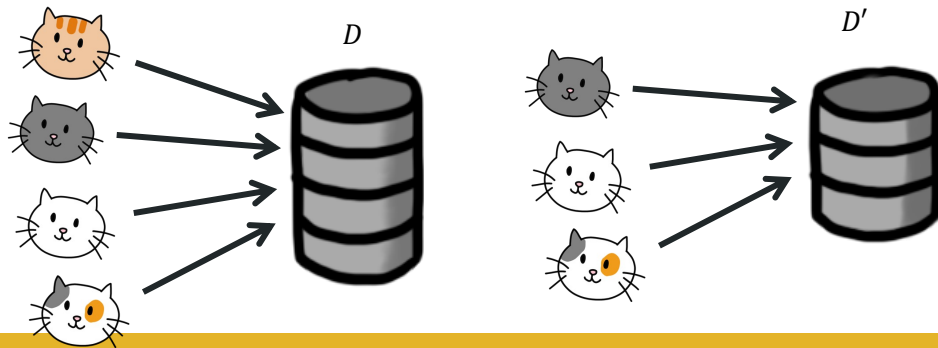
- Every cat else: 0 ~ 5 kg

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

Algorithm: You are allowed to make a query that returns the sum of the weights of the cats in the database.

Q: What is the ℓ_1 -sensitivity here?

A: $\Delta_1 = 5$



DP Mechanisms

- Multiple mechanisms provide Differential Privacy and can be applied to various systems.
- A few examples:
 - The Laplace Mechanism (DP, continuous outputs)
 - The Randomized Response Mechanism (DP, binary inputs/outputs)
 - General Discrete Mechanisms
 - The Exponential Mechanism (DP, discrete outputs)
 - The Gaussian Mechanism (approximate DP, continuous)

DP Mechanisms

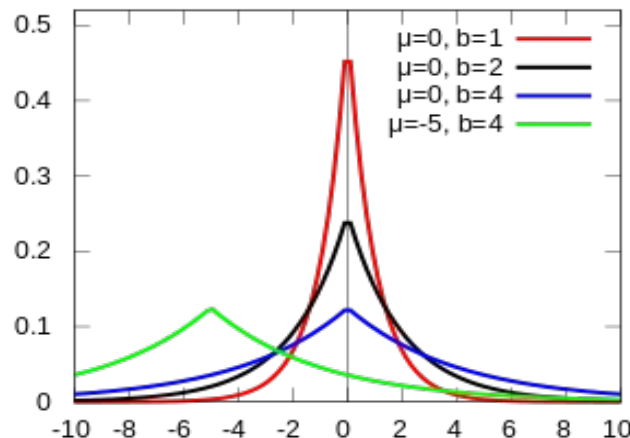
- Multiple mechanisms provide Differential Privacy and can be applied to various systems.
- A few examples:
 - **The Laplace Mechanism (DP, continuous outputs)**
 - The Randomized Response Mechanism (DP, binary inputs/outputs)
 - General Discrete Mechanisms
 - The Exponential Mechanism (DP, discrete outputs)
 - The Gaussian Mechanism (approximate DP, continuous)



https://en.wikipedia.org/wiki/File:Laplace,_Pierre-Simon,_marquis_de.jpg

Example: the Laplacian mechanism

- Let $Y \sim \text{Lap}(\mu, b)$
 - A Laplace distribution!
- With PDF: $p_Y(y) = \frac{1}{2b} e^{-\frac{|y-\mu|}{b}}$
- Usually, for DP, we set $\mu = 0$
 - So you may see $\text{Lap}(b)$ which is essentially $\text{Lap}(0, b)$
- $\text{Lap}(\mu, b)$ has variance $\sigma^2 = 2b^2$
- As b increases, the distribution becomes more flat



The Laplace Mechanism

- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -sensitivity of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

- Given any function f and its ℓ_1 -sensitivity, we can turn it into a DP mechanism if we add Laplacian noise to its output:

Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$ with ℓ_1 -sensitivity Δ_1 , the **Laplace mechanism** is defined as $M(D) = f(D) + (Y_1, Y_2, \dots, Y_k)$ where each Y_i is independently distributed following $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$.

The Laplace Mechanism

- Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$, and two neighboring datasets $D \in \mathcal{D}$ and $D' \in \mathcal{D}$, the ℓ_1 -sensitivity of f is the maximum change that replacing D for D' can cause in the output:

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

- Given any function f and its ℓ_1 -sensitivity, we can turn it into a differentially private mechanism if we add Laplacian noise to its output:

Given a function $f: \mathcal{D} \rightarrow \mathbb{R}^k$ with ℓ_1 -sensitivity Δ_1 , the Laplace mechanism is defined as $M(D) = f(D) + (Y_1, Y_2, \dots, Y_k)$ where each Y_i is independently distributed following $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$.

The Laplace mechanism provides ϵ -DP

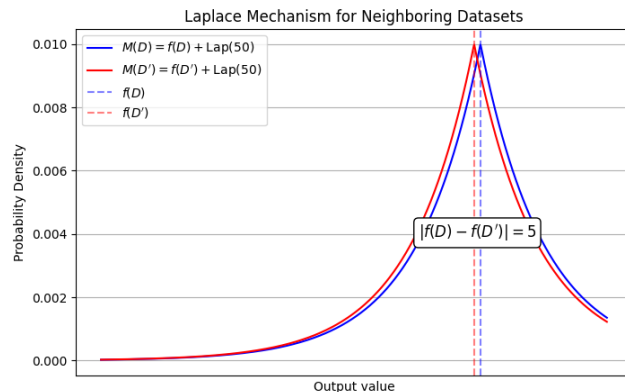
The Laplace Mechanism in our running example

- In our example: let's take $\epsilon = 0.1$, and together with $\Delta_1 = 5$, we have

$$M(D) = f(D) + \text{Lap}(b = \frac{\Delta_1}{\epsilon}) \Leftrightarrow$$

$$\Leftrightarrow M(D) = f(D) + \text{Lap}(\frac{5}{0.1}) \Leftrightarrow$$

$$\Leftrightarrow M(D) = f(D) + \text{Lap}(50)$$



The Laplace Mechanism in our running example

- $\epsilon = 0.1$ is good. What about $\epsilon = 4$?

You will not be affected, adversely or otherwise, by allowing your data to be used in any study or analysis, no matter what other studies, data sets, or information sources, are available. (bla bla... differential privacy ... bla bla). We have $\epsilon = 4$



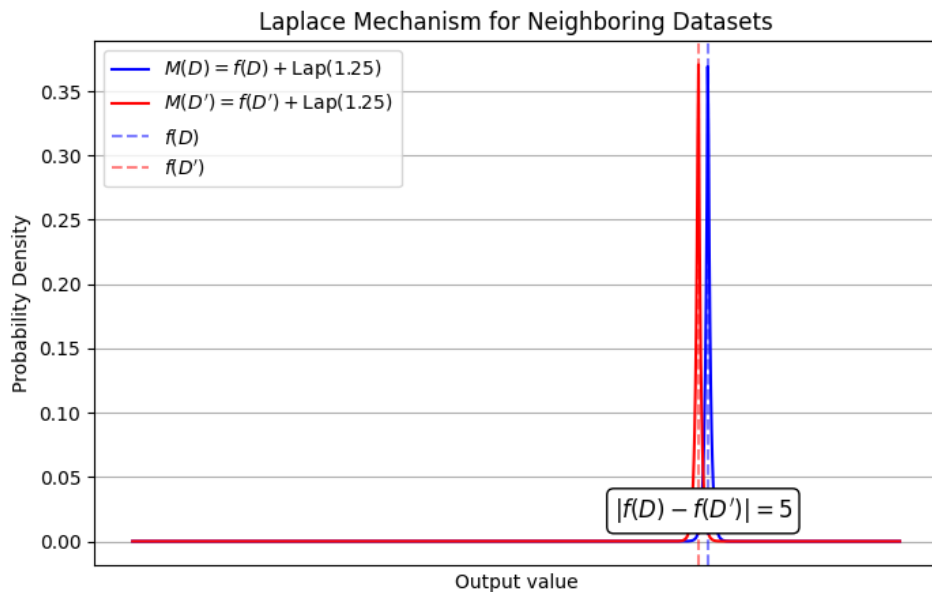
Is the ϵ small enough?

Should I share my data?



The Laplace Mechanism in our running example

- $\epsilon = 0.1$ is good. What about $\epsilon = 4$?



Awful!



A Few Other Nice Properties

Compositional privacy

- Given:

- $M_1 : D \rightarrow R_1$ being ϵ_1 -DP, and
- $M_2 : D \rightarrow R_2$ being ϵ_2 -DP

- This has a **gossip** analogy:

If A tells you something (potentially with noise), and then B tells you some other things (again, with noise), you may learn more by combining both pieces of information.

One may want to set a **total privacy loss budget** $\epsilon = \epsilon_1 + \epsilon_2 \dots + \epsilon_n$.

We can define a new mechanism:

$M : D \rightarrow R_1 \times R_2$ as $M(D) = (M_1(D), M_2(D))$.

Then, M is $(\epsilon_1 + \epsilon_2)$ -DP.

Group privacy

Theorem

Suppose mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is ϵ -differentially private. Suppose D, D' are two neighboring datasets $\in \mathcal{D}$ which differ in exactly k positions. Then:

$$\Pr(M(D) = R) \leq \Pr(M(D') = R) e^{k\epsilon}$$

- **Privacy guarantee drops linearly with the size of the group.**
- **TLDR:** If you need to hide the “effects” caused by a whole group of records, you need to prepare a larger privacy budget.

Approximate DP

- The following is a relaxation of the DP definition, that allows some tolerance:

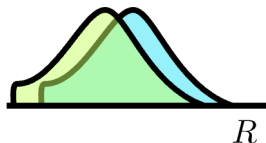
(Approximate) Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is (ϵ, δ) -differentially private $((\epsilon, \delta)$ -DP) if the following holds for all sets of possible outputs $S \subset \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

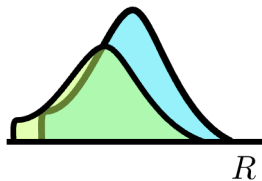
$$\Pr(M(D) \in S) \leq \Pr(M(D') \in S) e^\epsilon + \delta$$

- When $\delta = 0$, this is the same as ϵ -DP (called pure DP).
- What does this mean?

We have two distributions
 $f(R|D)$ vs $f(R|D')$



We multiply one
(e.g., blue) by e^ϵ



The area of the green one not covered by
the blue one now will be $\leq \delta$



Approximate DP: interpretation

(Approximate) Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is (ϵ, δ) -differentially private (ϵ, δ) -DP if the following holds for all sets of possible outputs $S \subset \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

$$\Pr(M(D) \in S) \leq \Pr(M(D') \in S) e^\epsilon + \delta$$

- A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ that provides ϵ -DP except for certain "bad" outcomes $B \subset \mathcal{R}$, where $\Pr(M(D) \in B) \leq \delta$ (for any $D \in \mathcal{D}$) also provides (ϵ, δ) -DP.
- This definition allows us to add less noise, if we are comfortable with the probability of bad outcomes

Approximate DP: interpretation

(Approximate) Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is (ϵ, δ) -differentially private (ϵ, δ) -DP) if the following holds for all *sets of possible outputs* $S \subset \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

$$\Pr(M(D) \in S) \leq \Pr(M(D') \in S) e^\epsilon + \delta$$

Theoretical distinction:

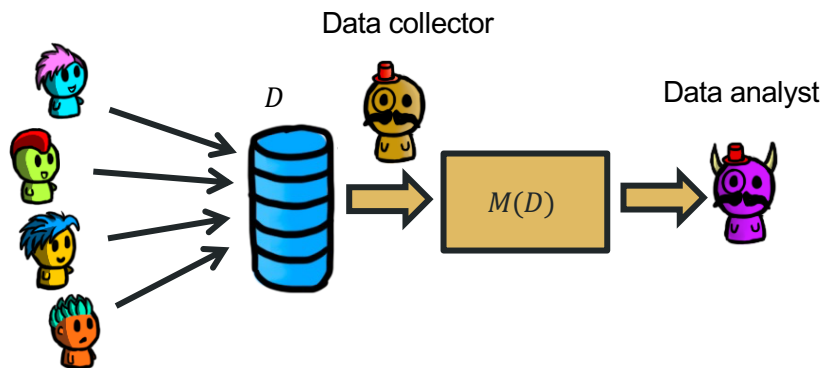
- $(\epsilon, 0)$ -differential privacy: the output observed is (almost) equally likely to be observed
- (ϵ, δ) -differential privacy: the output observed is much more or less likely to be observed

A Note on Differential Privacy Settings

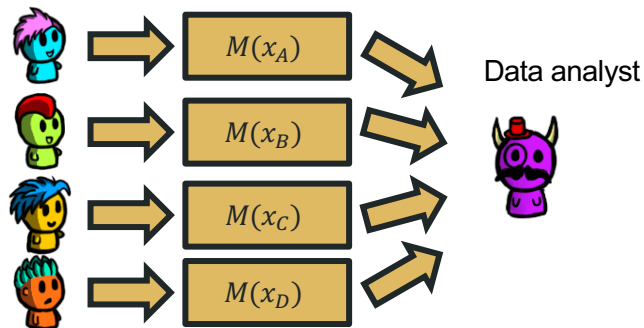
Central DP vs. Local DP

- Depending on who runs the mechanism, there are two broad models for differential privacy.

Central Differential Privacy: there is a centralized (trusted) aggregator



Local Differential Privacy: each user runs the mechanism themselves and reports the result to the adversary/analyst

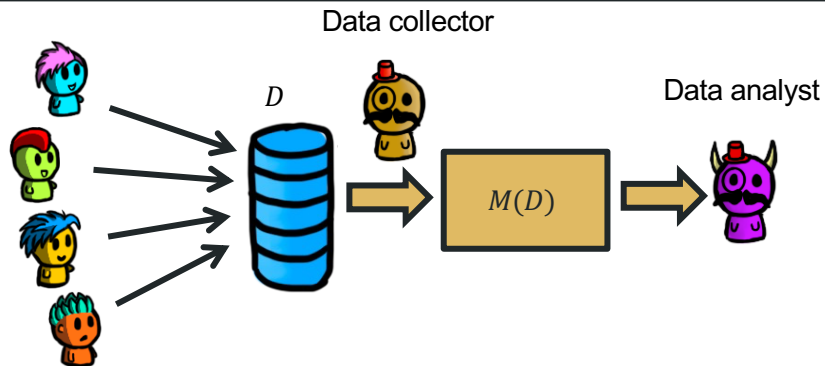


Central DP vs. Local DP

(Central) Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is ϵ -differentially private (ϵ -DP) if the following holds for all possible sets of outputs $R \subset \mathcal{R}$ and all pairs of neighboring datasets $D, D' \in \mathcal{D}$:

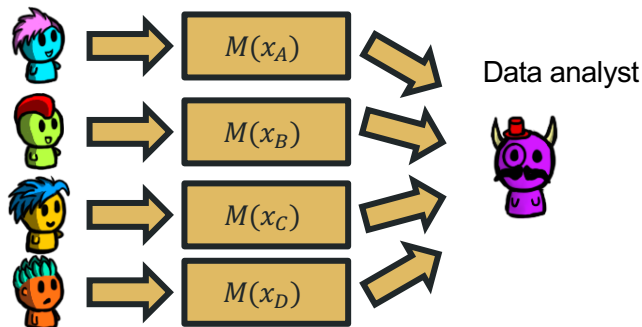
$$\Pr(M(D) \in R) \leq \Pr(M(D') \in R) e^\epsilon$$



(Local) Differential Privacy

A mechanism $M: \mathcal{D} \rightarrow \mathcal{R}$ is ϵ -differentially private (ϵ -DP) if the following holds for all possible sets of outputs $R \subset \mathcal{R}$ and all pairs of neighboring inputs $x, x' \in \mathcal{D}$:

$$\Pr(M(x) \in R) \leq \Pr(M(x') \in R) e^\epsilon$$



- They are “the same definition”, it’s just that the inputs to the mechanism and what we define as “neighbouring” inputs/datasets is usually different.

Central DP vs. Local DP

- Central DP

- Best accuracy, aggregation allows to hide in the crowd before we add noise.
- Need to trust the data collector.
- Hard to verify if noise was added.

- Local DP

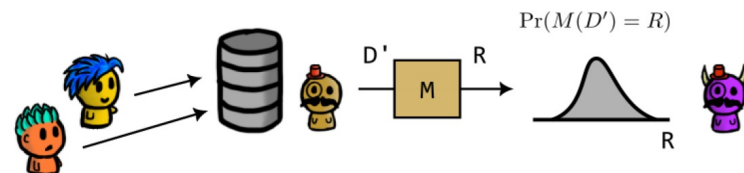
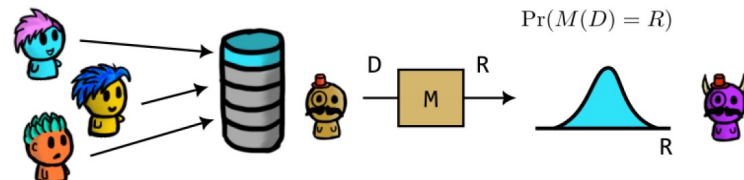
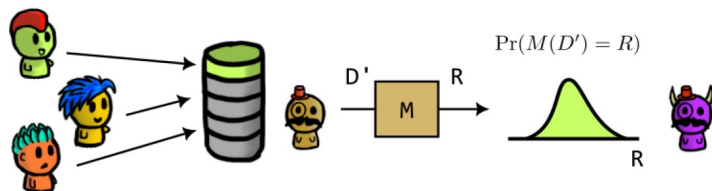
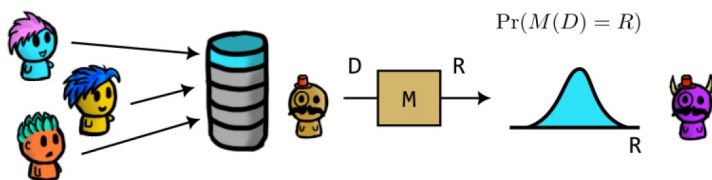
- Accuracy not as good. Each user adds noise which can compound in the final result.
- User doesn't need to trust anybody and knows they added noise.

Bounded DP vs. Unbounded DP

- There are two “main” definitions for how we define neighboring datasets in the central model.

Bounded DP: D and D' have the same number of entries but differ in the value of one.

Unbounded DP: D and D' are such that you get one by deleting an entry from the other one.



DP, but misused

The researchers' argument

- An example of large epsilon used in research:

Results: The misclassification rate of categorical variables ranged between 0.49 and 0.85 when the value of ϵ was 0.1, and it converged to 0 as ϵ increased. When ϵ was between 10^2 and 10^3 , the misclassification rate rapidly dropped to 0. Similarly, the mean squared error of the continuous variables decreased as ϵ increased. The performance of the model developed from perturbed data converged to that of the model developed from original data as ϵ increased. In particular, the accuracy of a random forest model developed from the original data was 0.801, and this value ranged from 0.757 to 0.81 when ϵ was 10^{-1} and 10^4 , respectively.

apply to the algorithm. In this study, a value of ϵ between 10^3 and 10^4 seemed heuristically appropriate; this depends on which data or model is used.



In this study, a value of ϵ between 10^3 and 10^4 seemed heuristically appropriate


Checkpoint on the Laplace Mechanism

(self-study)

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!




Q: what does smaller ϵ mean?

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!




Q: what does smaller ϵ mean?

A: more privacy

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!




Q: if we want more privacy, would we need to add more or less noise?

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!




Q: if we want more privacy, would we need to add more or less noise?

A: more noise. That's why $b \propto \frac{1}{\epsilon}$.

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!




Q: if changing D for D' can cause a huge change in $f(\cdot)$, is that a large or small sensitivity?

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!




Q: if changing D for D' can cause a huge change in $f(\cdot)$, is that a large or small sensitivity?

A: large sensitivity

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!




Q: if changing D for D' can have a huge impact in f , do we need a lot or a little noise to hide this impact?

The Laplace Mechanism – checkpoint!

The Laplace Mechanism: $M(D) = f(D) + Y$ where $Y \sim \text{Lap}(b)$ with $b = \frac{\Delta_1}{\epsilon}$ provides ϵ -DP

The variance is $2b^2$; higher b means more noise!



Q: if changing D for D' can have a huge impact in f , do we need a lot or a little noise to hide this impact?

A: a lot of noise.
That's why $b \propto \Delta_1$

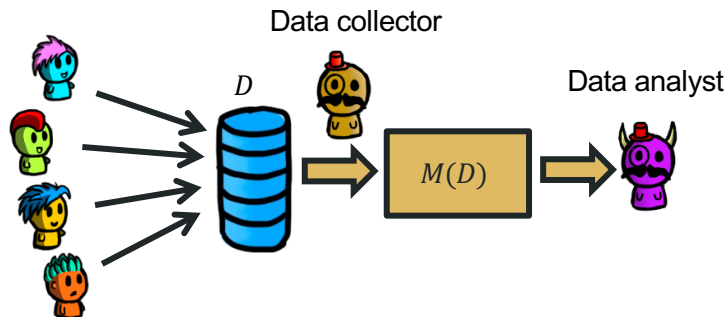
Laplace Mechanism: examples

Example 1: D contains the test results for virus X of a set of users. We want to release the total number of users that tested positive. How do we make this ϵ -DP?

- Under unbounded DP
- Under bounded DP

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



Laplace Mechanism: examples

Example 1: D contains the test results for virus X of a set of users. We want to release the total number of users that tested positive. How do we make this ϵ -DP?

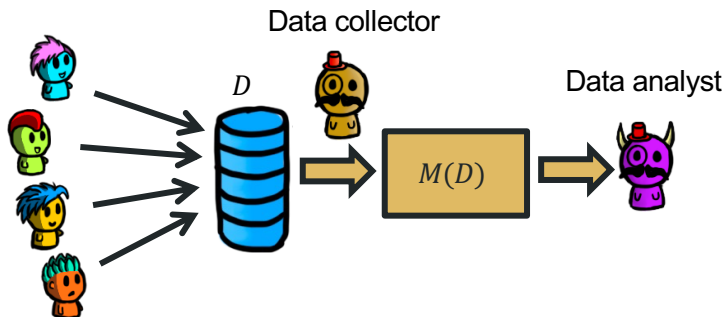
- Under unbounded DP
- Under bounded DP

A: sensitivity is 1 in both cases

Add $Y \sim \text{Lap}\left(\frac{1}{\epsilon}\right)$

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



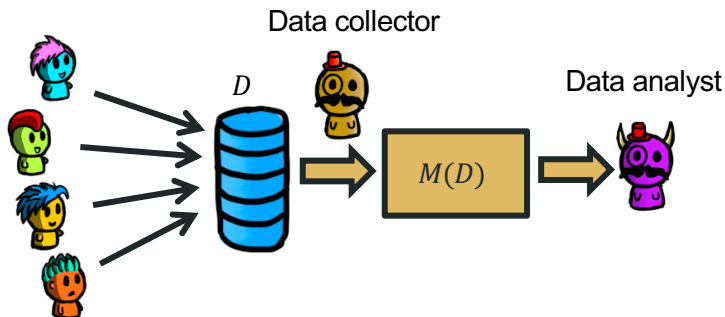
Laplace Mechanism: examples

Example 2: D contains the salaries of a set of users. The salaries range from 20k to 200k. We want to release the **total** salary of the users. How do we make this ϵ -DP?

- Under unbounded DP
- Under bounded DP

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



Laplace Mechanism: examples

Example 2: D contains the salaries of a set of users. The salaries range from 20k to 200k. We want to release the **total** salary of the users. How do we make this ϵ -DP?

- Under unbounded DP
- Under bounded DP

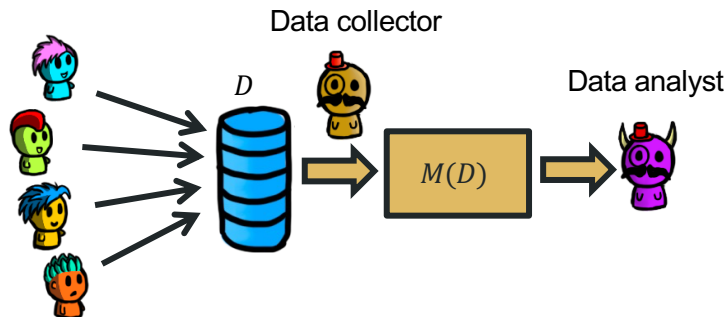
A: sensitivity is bounded by 180k in bounded DP and 200k in unbounded DP

Add $Y \sim \text{Lap}\left(\frac{180k}{\epsilon}\right)$ or

$$Y \sim \text{Lap}\left(\frac{200k}{\epsilon}\right)$$

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if } Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



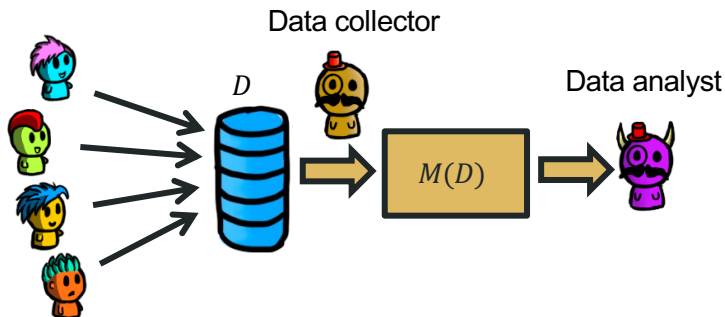
Laplace Mechanism: examples

Example 3: D contains the salaries of n users (n is public knowledge). The salaries range from 20k to 200k. We want to release the **average** salary of users. How do we make this ϵ -DP?

- Under bounded DP

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



Laplace Mechanism: examples

Example 3: D contains the salaries of n users (n is public knowledge). The salaries range from 20k to 200k. We want to release the **average** salary of users. How do we make this ϵ -DP?

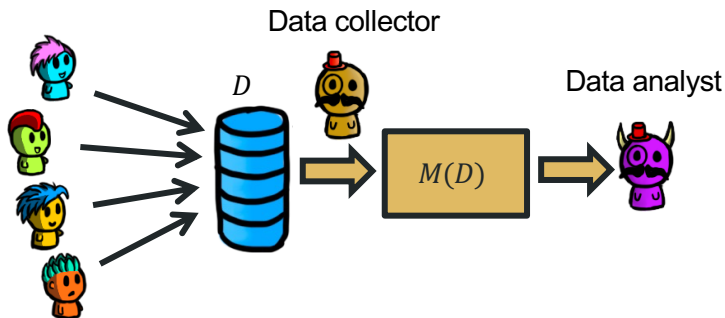
- Under bounded DP

A: sensitivity is bounded by $180k/n$

Add $Y \sim \text{Lap}\left(\frac{180k}{n\epsilon}\right)$

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



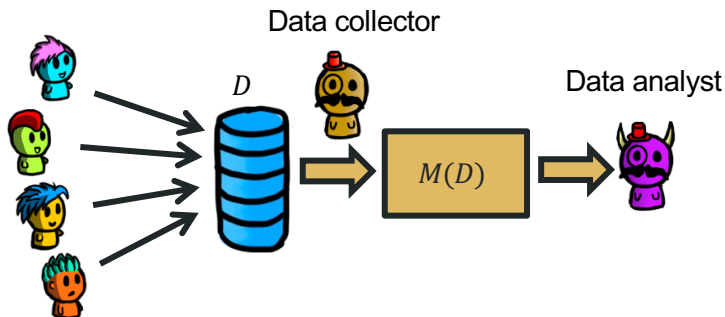
Laplace Mechanism: examples

Example 4: D contains the age of a set of users. We want to release the histogram of ages $[0-10)$, $[10-20)$... $[100,110)$. How do we make this ϵ -DP?

- Under unbounded DP
- Under bounded DP

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



Laplace Mechanism: examples

Example 4: D contains the age of a set of users. We want to release the histogram of ages $[0-10)$, $[10-20)$... $[100,110)$. How do we make this ϵ -DP?

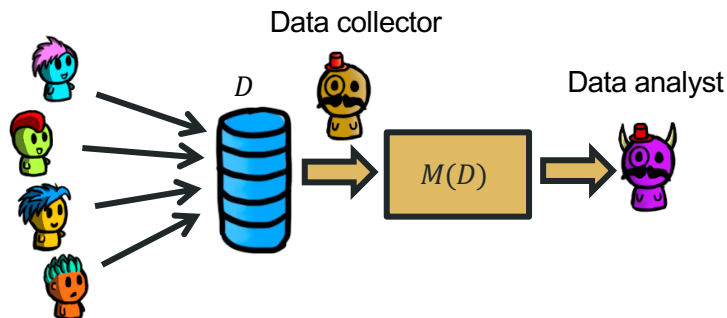
- Under unbounded DP
- Under bounded DP

A: sensitivity is 1 in unbounded 2 in bounded

Add $Y \sim \text{Lap}\left(\frac{1}{\epsilon}\right)$ or $Y \sim \text{Lap}\left(\frac{2}{\epsilon}\right)$ to each bucket in the histogram (drawn fresh for each bucket)

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if } Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$

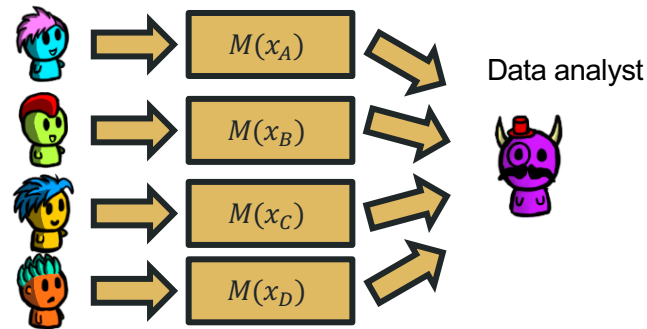


Laplace Mechanism: examples

Example 5: Alice wishes to report her age x_A in a differentially private way. It is public information that she is between 18 and 100 years old. She adds Laplacian noise with $b = 3$ to her age, and reports the resulting value. What is the level of DP that she gets?

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$



Laplace Mechanism: examples

Example 5: Alice wishes to report her age x_A in a differentially private way. It is public information that she is between 18 and 100 years old. She adds Laplacian noise with $b = 3$ to her age, and reports the resulting value. What is the level of DP that she gets?

A: sensitivity is bounded by 82

$$b = \frac{82}{\epsilon} = 3$$

$$\epsilon = 82/3$$

$$\Delta_1 \doteq \max_{D, D'} ||f(D) - f(D')||_1$$

$$f(D) + Y \text{ is } \epsilon\text{-DP if} \\ Y \sim \text{Lap}\left(\frac{\Delta_1}{\epsilon}\right)$$

