Differential Privacy in Practice

Enas Batarfi
Boston University

Overview

This report summarizes Assignment 3: Differential Privacy in Practice, which explores how data aggregation can still expose individual information and how differential privacy mitigates such risks.

We implemented a differentially private histogram using Laplace noise, analyzed the impact of different epsilon (ε) values on accuracy and privacy, and examined how repeated queries can lead to composition attacks that gradually reveal true values.

All code implementations, notebooks, and visualizations referenced in this report are included in the submission folder. The following sections present the methodology, results, and analysis for Questions 1-10.

Part 1: Plain Aggregates and Privacy

Question 1: Kinan's Music Preferences

First, looking at the count query, we can exclude any genre that does not appear for ages above 27, since I don't think Kinan could be younger than 27. That immediately rules out genres like Hip Hop, Country, and other entries that appear only in younger age groups.

This leaves a smaller set of possible options for Kinan: Pop, House, and Metal, since these appear for ages 27, 29, 31, and 32.

Next, I looked for public information about Kinan to narrow it down further. On X, I found a friend frequently mentioning him in posts about metal music. On his personal website, he referenced the metal band Vildhjarta, specifically their album "+ där skogen sjunger under evighetens granar +". Considering this evidence and Kinan's personality, Pop and House seem unlikely.

Combining the aggregate counts with these external clues, it strongly suggests that Kinan's preferred music genre is **Metal**. This example shows that even non-personally identifiable aggregate data can reveal private information when combined with public sources.

Question 2: Linking Attack and Age Inference

From Question 1, we assumed Kinan is older than 27, which lets us exclude most of the younger ages in the dataset. Looking at the dataset, we notice that most of the ages above 27 have only one record each. This means that just by looking at the counts, there are very few candidates for Kinan's data. The main possibilities were ages 29, 31, and 32.

To narrow it down, we performed a simple linking attack using public information. On **LinkedIn**, Kinan's education timeline shows:

- Bachelor's degree: 2012–2015
- Master's degree: 2016–2020
- PhD: completed in 2025

Assuming a typical academic path:

1. Start Bachelor's at 18 → finish at 22
2. Master's 23 → 27
3. PhD 27 → 31

This suggests Kinan is around **31 years old**. Next, we combine this with the findings from Question 1 about his music preference. In the dataset:

- Age 29 corresponds to house
- Age 31 corresponds to metal
- Age 32 corresponds to pop

Since Kinan's public references strongly indicate he likes metal, and pop or house do not fit, we can confidently infer that Kinan's data corresponds to **age 31 and metal**.

This linking attack demonstrates how starting with aggregate counts, then applying assumptions and external knowledge, allows us to refine the candidate records and confirm our previous findings from Question 1.

Question 3: Favorite Color

Looking at the age and color counts, most age groups have multiple colors, but for age 31 there is only one record and it corresponds to black. From previous analysis, we know Kinan is likely 31, so this points directly to black.

Even if Kinan were 32 instead, the record for age 32 also corresponds to black. This makes it obvious that Kinan's favorite color is black, regardless of the exact age within that range.

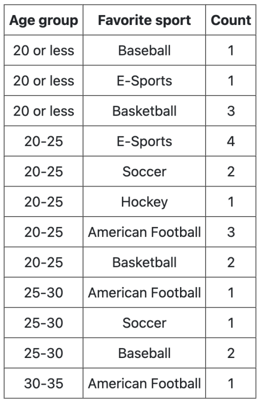
This inference is easy because the dataset is small and fine-grained, and the combination of age and color has very few records. Black also fits Kinan's musical taste and overall aesthetic. Thus, we can confidently conclude that Kinan's favorite color is **black.**

Question 4: Favorite Sport Reasoning

This question is trickier because the count is grouped by age ranges rather than exact ages. From our previous analysis, we know Kinan is likely 31, but he could also be 32. From the earlier queries, we saw that only two records in the dataset are above 30.

Looking at the current query, the 30–35 age group contains two different favorite sports, which indicates that each person in that age group has a different preference. This means that just from the current dataset, we cannot be completely certain about Kinan's favorite sport. The possibilities include any sport represented in that group.

To refine our guess, we then consider last year's data (see Fig. 1). Last year, Kinan would have been around 30 or 31, placing him in the last four rows of that dataset. By comparing the two datasets, we can identify which sports overlap for individuals in this age range. The only sport that consistently appears in both datasets for someone of Kinan's age is **baseball**.



Age group	Favorite sport	Count
20 or less	Baseball	1
20 or less	E-Sports	1
20 or less	Basketball	3
20-25	E-Sports	4
20-25	Soccer	2
20-25	Hockey	1
20-25	American Football	3
20-25	Basketball	2
25-30	American Football	1
25-30	Soccer	1
25-30	Baseball	2
30-35	American Football	1

Figure 1: Favorite sports from last year's dataset

Question 5: Differencing Attack

This process, in which we use differences across datasets to infer individual information, is known as a *differencing attack*. Combining this with what we know about Kinan's age and the dataset distribution helps narrow down the possibilities. Based on this reasoning, we can confidently conclude that Kinan's favorite sport is **baseball**, although there was some initial uncertainty due to the age grouping.

Part 2: Implementing Differential Privacy

Question 6: Effect of Varying Epsilon

Running `dp.py` with different ε values shows clear trends in the noise added to the histogram. Figure 2 compares three runs side by side.

- **Large ε (1, 5, 10):** Counts stay very close to the true values. The histogram looks accurate, but privacy is weaker because individual contributions are more obvious. For example, in Fig. 2, we can clearly see that the mode is at 1, the true value. With large ε, privacy protection is minimal.
- **Medium ε (0.5):** Noise is noticeable but moderate, balancing privacy and accuracy well. As we repeat the queries multiple times, the noisy counts start to cluster around the true value. For example, the most repeated value for the first row is around the real count of 1.
- **Small ε (0.1, 0.05, 0.01):** Noise dominates the counts, sometimes producing very high or negative values. Privacy is strong, but the histogram becomes less realistic. With so much noise, the accuracy of individual counts is greatly affected.

In summary:

- Large ε → weak privacy, counts close to true values.
- Medium ε (0.5) → decent privacy, mostly reasonable counts.
- Small ε → strong privacy, very noisy counts.

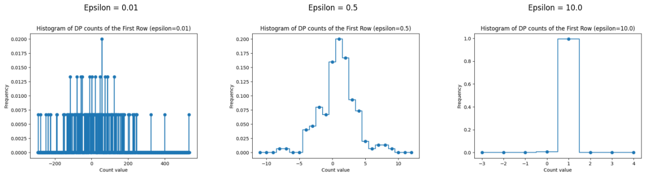


Figure 2: Histogram results for different ε values.

Question 7: Noised Histogram and Plot Analysis

For the first row in the histogram (age 19, Pop), the true count without noise is **1**. After running the differentially private histogram multiple times with ε = 0.5, the observed statistics, as illustrated in Figure 3, are:

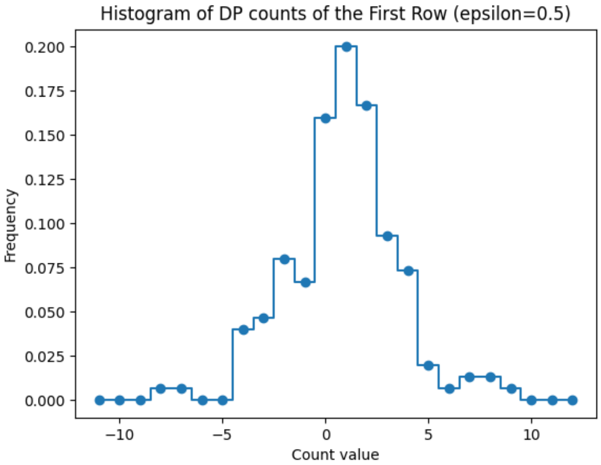


Figure 3: Histogram of observed counts for the first row over 150 runs with ε = 0.5.

- **Most likely value (mode):** 1
- **Median:** 1
- **Mean (average):** ∼1

Observations and interpretation:

- As seen in Figure 3, the mode, median, and mean all match the true count of 1, indicating that the Laplace noise did not significantly skew the observed values.
- Noise is noticeable but moderate. As we repeat the query multiple times, the noisy counts cluster around the true value; in this case, the most frequent value is exactly 1.
- Smaller ε values could introduce higher variability, increasing privacy but producing a wider spread of counts. While, larger ε values will reduce noise, keeping counts tightly clustered around the true value, which improves accuracy but weakens privacy.

Overall, this demonstrates that differential privacy introduces controlled randomness. With moderate ε values, the mechanism achieves a reasonable balance between protecting individual data and maintaining useful aggregate accuracy. However, repeated queries on the same dataset can eventually reveal true information, emphasizing the importance of carefully managing the privacy budget.

Part 3: Differential Privacy and Composition

Question 8: Composition Attack on Average Age

The exposed averages indicate that Kinan is most likely in either the **"5–8 Years"** or **"More than 10 Years"** programming experience group.

Considering that Kinan is **31 years old** and started his bachelor's degree in 2012, it is impossible for him to fall into any category with less than 5 years of experience. Therefore, the first two categories can be excluded.

Based on his educational timeline, it is highly plausible that Kinan has more than 10 years of programming experience. Even if he started programming in the last year of his bachelor's (2015), he would already have about 10 years of experience by 2025. However, the **"More than 10 Years"** category in the dataset includes only two participants with an average age of around 25–27, which makes it somewhat underrepresented for someone aged 31. For that to be true, the other participant would need to be around 20 years old with over 10 years of experience, implying they started programming before the age of 12, which is uncommon.

1. **Observations from the DP results:**
   - Each execution of the DP average query produces slightly different outputs due to Laplace noise.
   - Results for the last two experience categories typically range between 24–30, making precise identification challenging.
   - Patterns in the data combined with Kinan's profile indicate he is in one of these final two categories.
2. **Rationale for using the median:**
   - Both mean and median were tested across 200 and 2000 iterations.
   - The median provided more consistent results, particularly when outliers could distort the mean.
   - Both methods validated the accuracy for the count table and music genre table from Part 1.
3. **Approach for accuracy assurance:**
   - The composition attack was executed twice with 2000 iterations each to watch the differences.
   - Increasing iterations reduces privacy but improves the accuracy of the exposed results.
   - Repeated runs consistently placed Kinan in one of the final two categories, aligning with his age and academic timeline.

Question 9: Final Deduction

By running the DP count query grouped by programming experience, we were able to recover the noisy counts for each group. Although each run introduces noise, repeating the query multiple times reveals consistent patterns. Using the composition attack, we deduced the true counts used in the previous analysis which for the last two categories are:

- **5–8 Years:** 4 participants
- **More than 10 Years:** 2 participants

These counts support the logic behind the previously exposed averages:

- Kinan is **31 years old**, older than the averages of all groups except **“5–8 Years”** or **"More than 10 Years"**.
- The **“More than 10 Years”** group has an average age of 25–27 with only two members. For the average to hold, the other member would need to be around 20, which is unlikely given typical education timelines.
- The **"5–8 Years"** group has a higher average age and includes four members, which aligns well with Kinan being 31, making it statistically the most plausible group.
- Groups with less than 5 years of experience are impossible based on Kinan’s age and educational background.

Using the composition attack on both the programming level and the counts, we conclude with moderate confidence that Kinan could belong to either the **5-8 Years** group or the **More than 10 Years** group. Statistically, the first group appears more likely, but considering his educational timeline, it is plausible that he has more than 10 years of programming experience.

Attribute	Value
Age	31
Favorite Music	Metal
Favorite Sport	Baseball
Programming Experience	Most likely 5-8 years, possibly more

Table 1: Summary of Kinan’s personal information and inferred programming experience.

Question 10: Privacy Budget Enforcement

The `BudgetTracker` class does not actually prevent overuse of the privacy budget. It only tracks usage within the client code, which means that anyone could bypass it by directly calling the underlying functions or by creating a new `BudgetTracker` instance to reset the budget. Essentially, it is a simulation rather than a true enforcement mechanism.

A more robust approach would be to enforce the privacy budget at the data system or database level. The system should track the total privacy budget consumed and automatically block any queries once the budget is exhausted. Otherwise, composition attacks, such as those demonstrated in earlier questions could still leak sensitive information.