DP Wizard *Enhanced*: CS2080 Final Project

Isaac Lund, Brianna Chan, Yaying Liang Li, Shiloh Liu, Namat Noori

May 8, 2025

Introduction

Differential Privacy (DP) is an increasingly-common framework for computing statistics in a way that hides the presence or absence of any singular individual's data, yet effectively communicating DP understandings and implications to better support data analysts remains a pressing need. Various studies have been conducted surrounding communication strategies around differential privacy, and a list of interfaces containing textual, visualization, and interactive elements have been designed. However, many of these tools fail to clearly bridge technical definitions with real-world decisions users must make. Our goal is to close this gap by implementing new, concrete interactive features in DP Wizard that directly reflect findings from privacy communication literature.

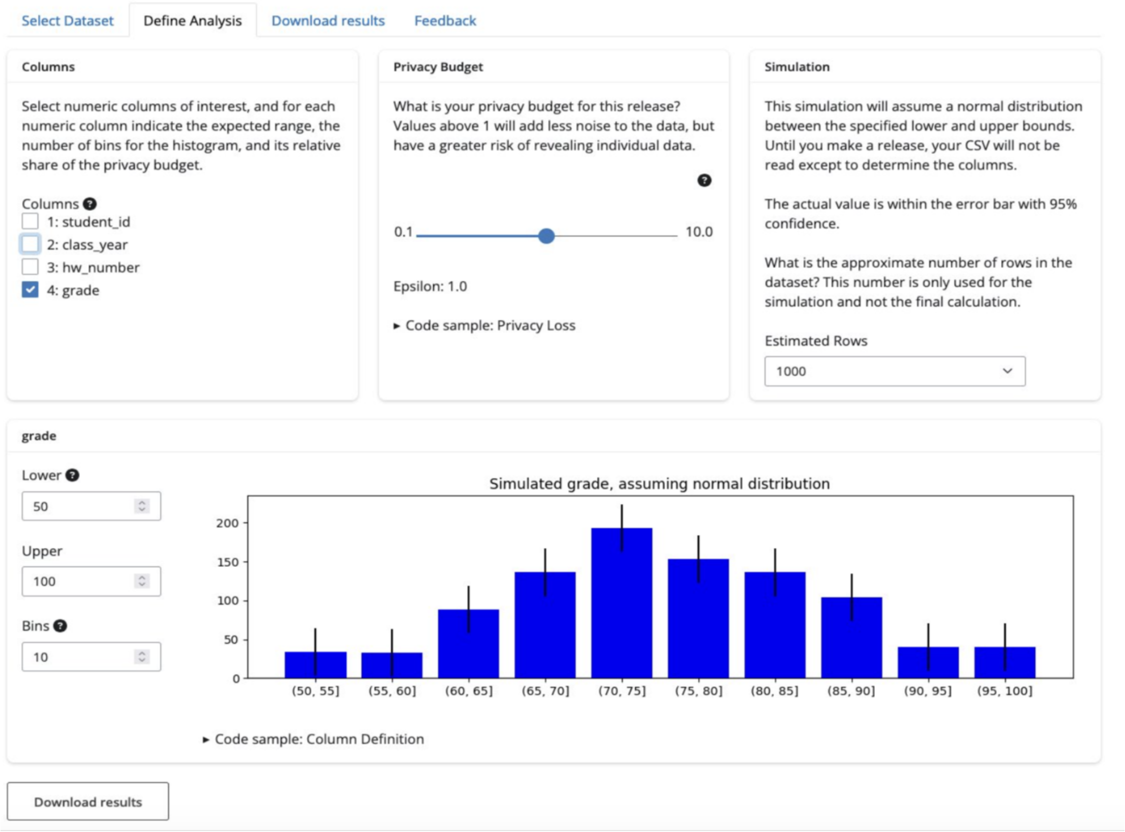
Our project design objectives build on the original DP Wizard goals: (1) to compute simple differentially private statistics without requiring deep expertise in data privacy, and (2) to help users understand how different ϵ values affect accuracy. DP Wizard Enhanced extends these foundations with two new goals: (1) enabling data analysts to understand what ϵ means in practice by visualizing the privacy-utility tradeoff and tracking ϵ budget usage during exploratory workflows, and (2) empowering analysts to perform realistic, privacy-preserving analysis. This includes support for a wider variety of differentially private query types, the use of filters, and iterative adjustment of parameters throughout the analysis process.

Background and Motivation

The Background and Motivation section will be organized as follows: (1) Communication Strategies around Differential Privacy, a general discussion of the key questions around communication; (2) Existing DP Interfaces, an overview of important highlights for existing interfaces that have been published; (3) Epsilon Visualization, a description of how experts have explained the parameter epsilon; (4) Privacy Budget Design, a description of how privacy budget allocation and management has been presented across interfaces.

Communication Strategies around Differential Privacy

We want to highlight several papers that emphasize the importance as well as challenges that come with the process of explaining the concept of differential privacy to non-privacy experts. Research suggests that trust is not rooted solely in mathematical guarantees but also in how DP is explained and whether those explanations align with users' concerns. Drawing on contextual integrity theory, users' willingness to share often hinges on what they think their data will be used



Columns

Select numeric columns of interest, and for each numeric column indicate the expected range, the number of bins for the histogram, and its relative share of the privacy budget.

Columns

□ 1: student_id
□ 2: class_year
□ 3: hw_number
☑ 4: grade

Privacy Budget

What is your privacy budget for this release? Values above 1 will add less noise to the data, but have a greater risk of revealing individual data.

0.1 ——●—— 10.0

Epsilon: 1.0

▸ Code sample: Privacy Loss

Simulation

This simulation will assume a normal distribution between the specified lower and upper bounds. Until you make a release, your CSV will not be read except to determine the columns.

The actual value is within the error bar with 95% confidence.

What is the approximate number of rows in the dataset? This number is only used for the simulation and not the final calculation.

Estimated Rows: 1000

grade

Lower: 50

Upper: 100

Bins: 10

▸ Code sample: Column Definition

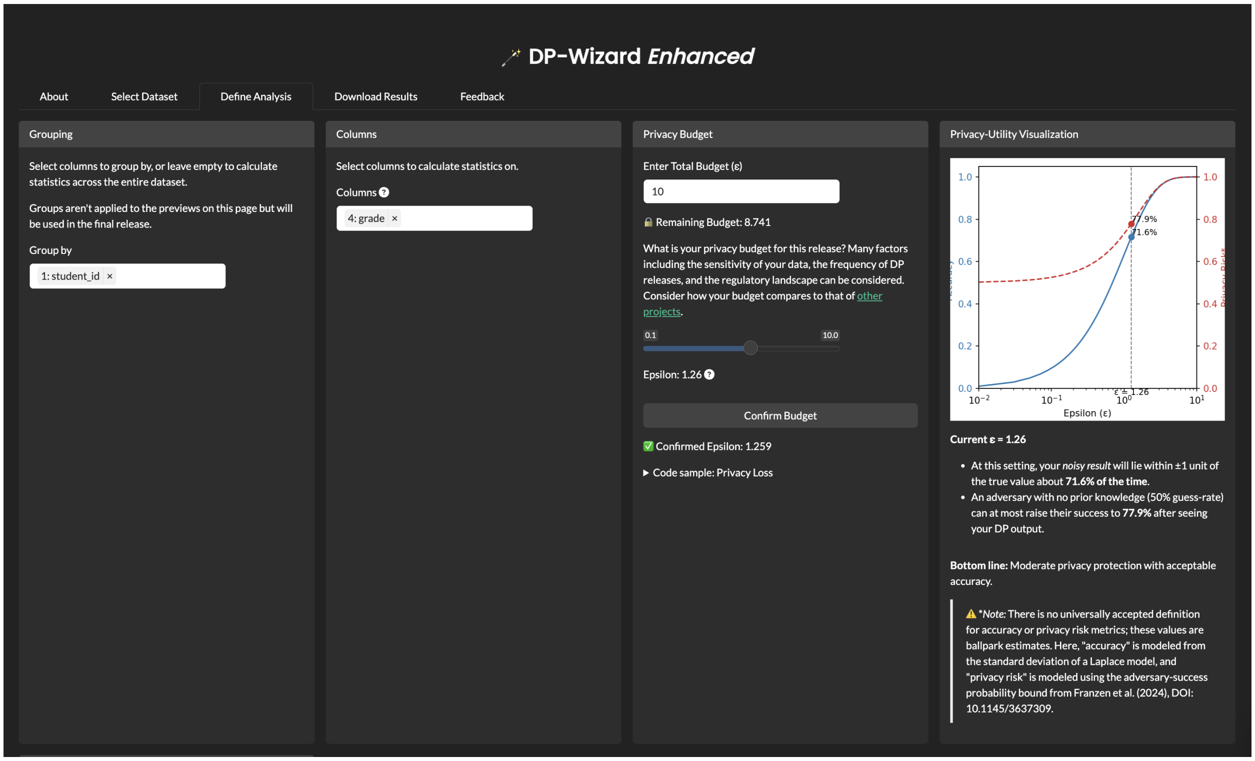
Figure 1: Overview of Original DP Wizard Interface.

for — for instance, whether it is shared with third parties or simply aggregated into a public graph. Thus, communication strategies around DP must be attuned to context-specific expectations.

DP Wizard currently serves data owners and analysts—individuals who are actively applying privacy mechanisms to datasets, yet the software currently lacks certain features that are pivotal for real, secure data analysis. To better understand its strengths and limitations, we spoke with one of DP Wizard's developers, Chuck, and reviewed open issues and feedback discussions on GitHub. We also spoke with one of OpenDP's developers, Michael, who significantly contributed to changes to the standard deviation query. These conversations highlighted a few limitations in the current interface, such as limited query types, lack of support for budget tracking during iterative exploration, and insufficient clarity on how parameter changes affect privacy guarantees. Our redesign builds on this feedback directly. We aim to extend the usability of DP Wizard for iterative exploratory data analysis by offering support for more realistic query workflows, making the usage of ϵ more transparent and improving the interpretability of results for analysts.

These improvements also reflect key principles in the literature on usable privacy tools. Prior work has emphasized the importance of aligning explanations with users' mental models, the need for tools to support iteration and budget-awareness, and the necessity of reducing cognitive load in interfaces meant for non-experts, which we incorporated into our redesign to ensure analysts not only achieve compliance, but feel confident and informed throughout their workflow.

Furthermore, studies have described how personal attributes influence subjects' willingness to



Select columns to group by, or leave empty to calculate statistics across the entire dataset.

Groups aren't applied to the previews on this page but will be used in the final release.

Group by: 1: student_id ×

Select columns to calculate statistics on.

Columns: 4: grade ×

Enter Total Budget (ε): 10

🔒 Remaining Budget: 8.741

What is your privacy budget for this release? Many factors including the sensitivity of your data, the frequency of DP releases, and the regulatory landscape can be considered. Consider how your budget compares to that of other projects.

0.1 — 10.0

Epsilon: 1.26

✅ Confirmed Epsilon: 1.259

▶ Code sample: Privacy Loss

Current ε = 1.26

- At this setting, your *noisy result* will lie within ±1 unit of the true value about **71.6% of the time.**
- An adversary with no prior knowledge (50% guess-rate) can at most raise their success to **77.9%** after seeing your DP output.

Bottom line: Moderate privacy protection with acceptable accuracy.

> ⚠ **Note:* There is no universally accepted definition for accuracy or privacy risk metrics; these values are ballpark estimates. Here, "accuracy" is modeled from the standard deviation of a Laplace model, and "privacy risk" is modeled using the adversary-success probability bound from Franzen et al. (2024), DOI: 10.1145/3637309.

Figure 2: Overview of DP Wizard Enhanced Interface.

share data and engage in informed decision-making, and these concepts are highly transferable to the experience of data analysts and developers who must interpret, implement, and justify differential privacy mechanisms. While they may not be the original data contributors, data analysts are still end-users of DP systems — they must make decisions about privacy budgets, choose ϵ values, and communicate the implications of those choices to downstream stakeholders. If the interface overpromises privacy protections or obscures the trade-offs involved, it risks becoming another form of privacy theater — one where developers feel secure but are operating under misunderstandings, as expanded upon in Smart et al. (2022) "Understanding Risks of Privacy Theater with Differential Privacy."

Cummings et al. (2021) conducted one of the first large-scale investigations into user expectations for differential privacy. They surveyed over 2,400 participants to assess whether people care about the protections DP provides and if knowledge of DP increases data-sharing willingness. They found that users do value DP, especially when it is presented in terms that resonate with their own privacy concerns. However, the way DP is described greatly influences expectations. Vague or over-technical descriptions can lead to "privacy theater"—misplaced confidence in protections. To avoid this, descriptions must match what users care most about, providing transparency around how and why their data is protected. The authors categorize these different descriptions into six main classifications. They find that the "Trust," "Risk," "Enables," and "Control" categories increase the likelihood that respondents think their information would be secured against disclosures to a data analyst, as well as the likelihood that respondents think their information would not be disclosed through graphs or charts made using their information. The paper demonstrates that users' mental models of privacy protection are shaped by explanations that match their underlying concerns, such as individual disclosure or use in visualizations. To align with this, for example, we designed our epsilon visualizations to convey disclosure risk clearly and interactively, reinforcing trust by directly targeting one of the most important user concerns: the risk of being personally identified.

More broadly, Dibia et al. (2024) synthesized these themes across user studies and interface evaluations. Particularly in Table 5: Summary of challenges and recommendations for DP software tools, the authors address the need for more specificity and support surrounding pillars including target expertise, budgeting, utility analyses, and correctness checking. These pillars clearly align with the goals of our own project. Inspired by Dibia's recommendations, we structured our interface to minimize users' cognitive load by guiding them through each choice with tailored visuals and explanations. Instead of asking users—especially developers who may not have deep familiarity with ϵ, δ, or sensitivity settings—to manually configure privacy parameters upfront, we focus on first communicating what these values mean through visual and interactive elements. Additionally, since the privacy budget quantifies how much privacy loss is allowed when analyzing or releasing data, it plays a central role in any practical DP deployment. Currently, there exists no clear agreed-upon value for privacy budget, and users may have difficulty setting and interpreting privacy budget. It is especially important to create an interface that allows users to adjust the privacy budget based on their own data which they would upload to the website. Our work to visualize and control the use of the privacy budget is therefore foundational. The authors also suggest to "always raise errors for when DP might be violated," and we implemented this in DP Wizard, Enhanced by changing the histogram stylistically when DP is violated and the privacy budget is exceeded.

Existing DP Interfaces

In response to usability concerns as illustrated in the previous section, a range of graphical and interactive interfaces have emerged. These tools aim to make differential privacy more understandable, operationalizable, and responsive to users' needs. We reviewed several prior DP tools to understand which design patterns and educational scaffolds have successfully improved comprehension or decision-making for non-experts. These tools helped shape the rationale for our enhancements to DP Wizard.

PSI (Gaboardi et al. 2016) is an early example targeting data depositors and analysts. PSI abstracts DP complexity through simple query interfaces and strong default protections. PSI seeks to call attention to three different scenarios in which differential privacy can provide a clear benefit over other approaches toward privacy. The authors want PSI to support the idea that differentially private statistics can often be indistinguishable from non-private analyses, especially in samples where n is large, meaning that differential privacy sacrifices little utility. Secondly, PSI should broaden access to sensitive data because incorporating differential privacy protects such data and renders it usable. PSI is particularly useful in regards to the third use case, where differential privacy is able to provide meaningful statistical information in a scenario where any statistical information is better than no information at all. In cases where data is unavailable due to restrictive and time-consuming provisions, these metrics become especially valuable.

Building on this insight, our project expands DP Wizard's capabilities by adding statistical query types that reflect real-world analyst needs and reinforce the utility-based justifications seen in PSI. While the original DP Wizard supported three query types (Histogram, Mean, and Median), we implemented three additional functions: Count, Quantile, and Standard Deviation. These additions directly reflect the utility-focused spirit of PSI, aiming to make differentially private tools more responsive to the real-world analytical needs of researchers. Count queries help capture basic tabulations essential for almost any dataset, while quantile and standard deviation queries are key components of understanding data spread and statistical confidence. In contexts where differentially private insights must substitute for unavailable raw data, these additions ensure that analysts can still compute foundational summary statistics under meaningful privacy guarantees. Through these enhancements, we continue the PSI vision—supporting useful, privacy-safe analysis where traditional access barriers might otherwise prevent it.

DP Creator (Sarathy et al. 2023) extended these ideas for developers as a later iteration and prototype of the original PSI. It tracked cumulative ϵ consumption across multiple queries and included interface elements (sliders, progress bars) for tuning and budgeting. DP Creator provides an end-to-end environment where a user can load data, choose analysis queries (like counts, sums, etc.), and then export results with DP applied. Importantly, it keeps track of the cumulative privacy budget – if multiple queries are asked, the tool ensures the total ϵ remains within a user-specified limit.

Sarathy et al.'s user study with 19 data practitioners found that even experienced analysts required better explanations and guidance of what ϵ means. This is especially challenging since analysts also had to think carefully about which queries were worth using up some of the privacy budget, a new decision not present in standard analysis. Drawing directly from DP Creator's design, which includes a visual privacy budget tracker and adjustable ϵ sliders, we implemented a budgeting panel in DP Wizard Enhanced where users can preview the ϵ cost of a query and adjust it interactively via a slider. We also incorporated a confirmation step before query submission, supported by real-time visualizations of disclosure risk, to help users assess whether the privacy–utility trade-off is acceptable. While our current implementation focuses on single-query workflows, we plan to extend this in future iterations to support cumulative ϵ tracking across multiple queries.

Other interfaces include DPComp (Hay et al. 2016), which enables side-by-side comparisons of DP algorithms. Comparison of different differential privacy algorithms and their various privacy guarantees can be a fruitful direction to explore in the future as we implement more functionalities for DP Wizard. On the other hand, the DPP (St. John et al. 2021) was also introduced, and it guides users (especially lay data owners) through parameter selection by translating ϵ into more relatable outcomes, like "risk of guessing someone's data." With minimal instruction, these data owners were able to use the tool's visual feedback to set an ϵ that achieved an acceptable balance of privacy and accuracy for their scenario. We were inspired by this finding to go beyond visualizations alone and include a textual summary that distills the privacy–utility tradeoff into a simple, decision-oriented takeaway. Since DPP shows that users benefit from plain-language explanations that connect ϵ to real-world consequences in human-centered terms, our DP Wizard Enhanced interface also seeks to include bullet-pointed explanations and a bottom-line summary to translate ϵ into interpretable outcomes.

Communicating Epsilon Budget and Tradeoffs

Franzen et al. (2024) explicitly addressed how to support informed privacy decision-making by communicating the privacy–utility tradeoff. One key feature was the use of icon arrays to visualize incremental risk. For example, their design shows that moving from $\epsilon = 0$ to $\epsilon = 0.1$ might only increase risk by 2% when the base risk is 20%, but doing so will increase the risk by 9% when the base risk is 90%. Visualization techniques like those used in ViP (Nanayakkara et al. 2022) and Panavas et al. (2024) also provide valuable guidance. ViP used quantile dot plots, derived from inverse CDFs, and confidence intervals to help users visually understand how ϵ affects output variance and risk. Panavas et al. expanded on this by building a modular interface that included ϵ, δ, sensitivity, and composition.

Besides communicating the effect of epsilon on privacy-utility tradeoffs, there is also a need for communicating epsilon's impact on analyses' privacy budgets. In exploratory settings, interfaces such as Overlook (Thaker et al. 2022) and the Measure-Observe-Remeasure (MOR) paradigm (Nanayakkara et al. 2024) reflect how DP needs to support iteration. For instance, MOR zeroes in on users' general struggle at grasping the concepts of randomness and probability distributions by introducing the idea of incremental ϵ spending. Instead of committing all budget upfront, users could "remeasure" queries they care about. This aligns with real-world data exploration, where insights emerge through trial and error. Thus, we also oriented our enhanced DP Wizard to explicitly support this needed iterative process.

Problem Definition

DP Wizard is an open-source web-based tool designed to help users—particularly data analysts and developers—interactively generate differentially private statistics from uploaded datasets. It was originally built to make differential privacy concepts more accessible to non-experts by helping users explore how privacy parameters (such as ϵ) impact their data analysis tasks. After selecting a local CSV file, users are prompted to describe the analysis they need by selecting columns to group by and calculate statistics on. They also select a privacy budget for the release. The system then provides output statistics with differentially private guarantees, while visualizing how privacy loss is distributed across queries.

However, the original version of DP Wizard had limited functionality and minimal feedback to help users understand the privacy implications of their actions. To address these limitations, DP Wizard Enhanced extends the interface with targeted features that align closely with our core design objectives. These improvements aim to make the tool both more useful for actual data analysis use cases and more informative for non data-privacy experts as they approach understanding differential privacy:

- Added support for Count, Quantile and Standard Deviation queries, expanding the analyst's ability to compute a wider range of summary statistics under DP constraints. These are essential statistics in nearly any exploratory analysis, and their inclusion aligns with PSI's emphasis on utility-preserving tools. These additions also reflect our goal of empowering analysts to perform realistic analyses and making differentially private analyses accessible.

- A dynamic budget tracking system that provides real-time warnings when the privacy budget is low or exceeded. This is implemented through visual cues and alerts based on query-specific consumption and the user-defined value ϵ. To help analysts manage cumulative privacy loss over time, DP Wizard Enhanced thus achieves our objective of making ϵ usage more transparent and actionable.

- A visualization of how ϵ affects accuracy and privacy risk in real time. As users adjust ϵ, the graph dynamically updates to display two curves: one representing the expected accuracy of results, and the other showing the increased success probability of a potential adversary.

This visualization, along with the textual summaries and bottom line that help the user interpret quantitative percentages into qualitative descriptions of privacy protection, directly help reduce the user's cognitive load.

These changes aim to reduce user confusion, support iterative analysis, and better communicate trade-offs inherent in privacy-utility decision making. We hope that these designs enhance DP Wizard so that it is both a practical and meaningful interface for non data-privacy experts. For more information, please visit our *Github*.

Methodology

DP Statistics

In addition to the 3 queries already accessible through DP-Wizard (Histogram, Mean, Median), we added 3 more query types to the interface: Count, Quantile, and Standard Deviation. These are essential statistical tools that significantly expand the utility of the interface, especially for users performing exploratory or confidence interval analysis.

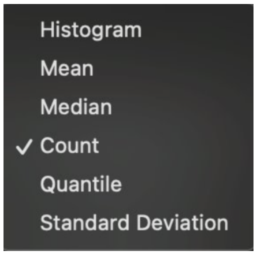
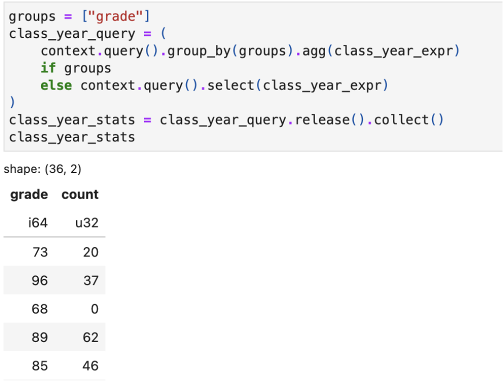


Figure 3: Dropdown Menu.

These functionalities were accomplished by creating separate folders for each query type, using OpenDP's Transformation libraries and Polars' API functionality, and linking these functions to the backend. We then updated the frontend so users could easily select these options from the dropdown interface during analysis.

DP Count

Count queries are simple yet foundational for initial data exploration and other query types, so implementing this query type would provide analysts with benefit. For example, knowing the count of the dataset gives us an immediate sense of scale of the dataset and helps us with filtering (ie. do we have enough observations in a subgroup to draw reliable conclusions?) Count queries are also beneficial when data analysts are calculating proportions (ie. when we're calculating mean or standard deviation and we need to find the denominator). Here is what the count query output looks like after being downloaded:



```
groups = ["grade"]
class_year_query = (
    context.query().group_by(groups).agg(class_year_expr)
    if groups
    else context.query().select(class_year_expr)
)
class_year_stats = class_year_query.release().collect()
class_year_stats
```

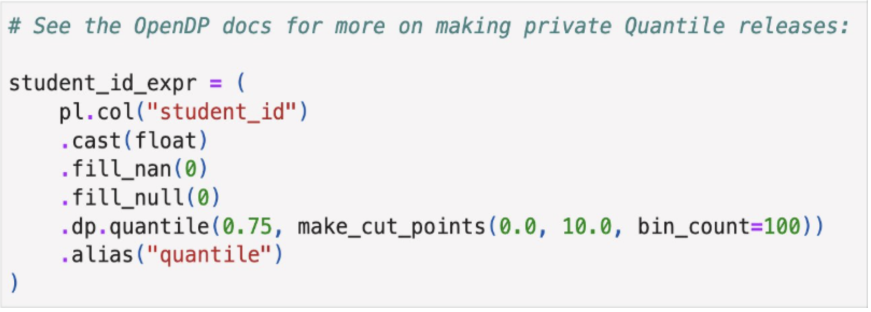
shape: (36, 2)

grade	count
i64	u32
73	20
96	37
68	0
89	62
85	46

Figure 4: DP Count Query Output Example.

DP Quantile

Quantile queries provide data analysts with distribution shape insights. The original DP-Wizard provides median (in other words, the 50th percentile) information, but our enhanced version provides information on the 25th or 75th percentiles as well. In other words, users are more easily able to see where the ”bulk” of the data lies and also how skewed it is. Here is what the quantile query looks like.



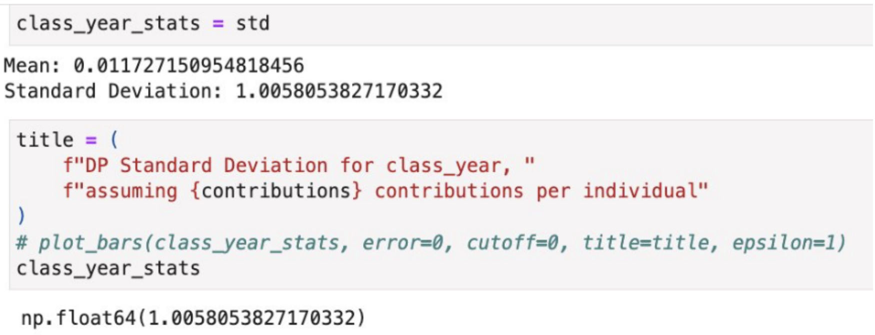
```
# See the OpenDP docs for more on making private Quantile releases:

student_id_expr = (
    pl.col("student_id")
    .cast(float)
    .fill_nan(0)
    .fill_null(0)
    .dp.quantile(0.75, make_cut_points(0.0, 10.0, bin_count=100))
    .alias("quantile")
)
```

Figure 5: DP Quantile Query Output Example.

DP Standard Deviation

Standard deviation queries are beneficial to data analysts because they generally quantify data points’ variability from the mean. This helps analysts interpret fluctuations in the data set and build private confidence intervals. Here is what the standard deviation query looks like after being downloaded.



```
class_year_stats = std
```

```
Mean: 0.011727150954818456
Standard Deviation: 1.0058053827170332
```

```
title = (
    f"DP Standard Deviation for class_year, "
    f"assuming {contributions} contributions per individual"
)
# plot_bars(class_year_stats, error=0, cutoff=0, title=title, epsilon=1)
class_year_stats
```

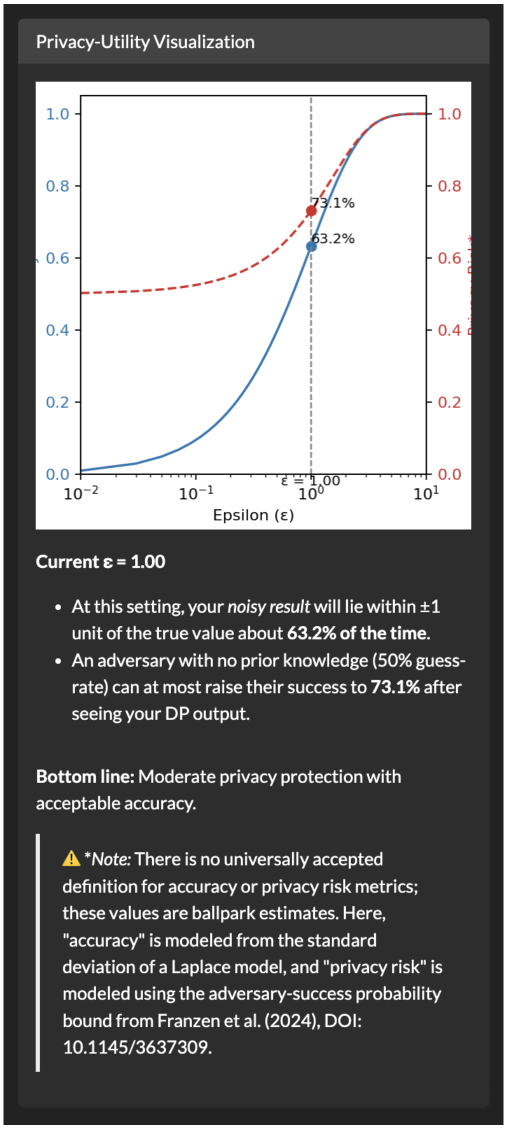
```
np.float64(1.0058053827170332)
```

Figure 6: DP Standard Deviation Query Output Example.

However, standard deviation statistics aren't supported by OpenDP as of May 10th, 2025 (confirmed by OpenDP developer Michael), so we had to manually implement the differentially private release of standard deviation statistics. This was done by calculating the mean along with some count queries, then finding the square root of the variance (see Github for more details).

Epsilon Visualization

To bridge the gap between abstract differential privacy parameters and tangible user understanding, we developed a real-time epsilon visualization panel. In our implementation, the core of our explanation of the privacy–utility tradeoff is a two-curve plot showing Accuracy (on the left y–axis) and Privacy Risk (on the right y–axis) adjacent to the epsilon slider:



Current ε = 1.00

- At this setting, your *noisy result* will lie within ±1 unit of the true value about **63.2% of the time.**
- An adversary with no prior knowledge (50% guess-rate) can at most raise their success to **73.1%** after seeing your DP output.

Bottom line: Moderate privacy protection with acceptable accuracy.

> ⚠ **Note:* There is no universally accepted definition for accuracy or privacy risk metrics; these values are ballpark estimates. Here, "accuracy" is modeled from the standard deviation of a Laplace model, and "privacy risk" is modeled using the adversary-success probability bound from Franzen et al. (2024), DOI: 10.1145/3637309.

Figure 7: Epsilon Visualization

For the user, there is a vertical dashed line that denotes the specific ϵ selected via the slider. Moreover, the color-coded curves—blue for accuracy and red for risk—intuitively indicate the trade-off: "if I move the slider to the right, my answer gets more accurate—but I also leak more privacy." This interactive, color-annotated design was chosen to reduce cognitive load and reinforce learning without requiring a background in DP. Instead of novice users needing to learn complex math behind Laplace noise or ϵ-differential privacy, they can simply adjust the vertical line and choose the risk–accuracy balance they are comfortable with. The two functions being graphed are

$$\text{Accuracy}(\varepsilon) = 1 - e^{-\varepsilon} \quad \text{and} \quad \Pr(D \mid y) = \frac{e^{\varepsilon}\, p_0}{e^{\varepsilon}\, p_0 + (1 - p_0)},$$

for Accuracy and Privacy Risk, respectively. However, below the visualization in the program is a note that concedes to users that these are simply ballpark figures for accuracy and privacy rather than concrete guarantees. The graphs were derived in the following ways:

Starting with the Accuracy curve, in class we defined the Laplace mechanism by drawing noise, with scale b, from the distribution

$$X \sim \mathrm{Lap}(b),$$

whose density is

$$f(y) = \frac{e^{-|y|/b}}{2b}.$$

Because this noise is centered at zero, the *accuracy* of a release (i.e. the probability that the noise shifts the answer by at most ±1) is

$$\Pr\big(|X| \le 1\big) = \int_{-1}^{1} \frac{e^{-|y|/b}}{2b} dy = 1 - e^{-1/b}.$$

Substituting $b = 1/\varepsilon$ to satisfy ε-DP yields

$$\mathrm{Accuracy}(\varepsilon) = 1 - e^{-\varepsilon}.$$

For the Privacy–Risk curve, we based our figure on Franzen et al. (2024)'s methodology for calculating privacy risk. Namely, we frame the adversary's task as a hypothesis test: did the mechanism run on dataset D (including the target individual) or on its neighbor D' (excluding them)? As done in the paper, we assign a baseline prior $\Pr(D) = p_0 = 0.5$ because an adversary has an equal likelihood of guessing between the two possible outcomes, randomly. Finally, to turn these understandings into a curve, we assign likelihoods $\Pr(y \mid D)$ and $\Pr(y \mid D')$, and then Bayes' Rule gives

$$\Pr(D \mid y) = \frac{\Pr(y \mid D)\, p_0}{\Pr(y \mid D)\, p_0 + \Pr(y \mid D')\, (1 - p_0)}.$$

From here, we know that differential privacy guarantees

$$\frac{\Pr(y \mid D)}{\Pr(y \mid D')} \le e^{\varepsilon},$$

so in the worst case $\Pr(y \mid D) = e^{\varepsilon} \Pr(y \mid D')$. Substituting and canceling the common factor yields

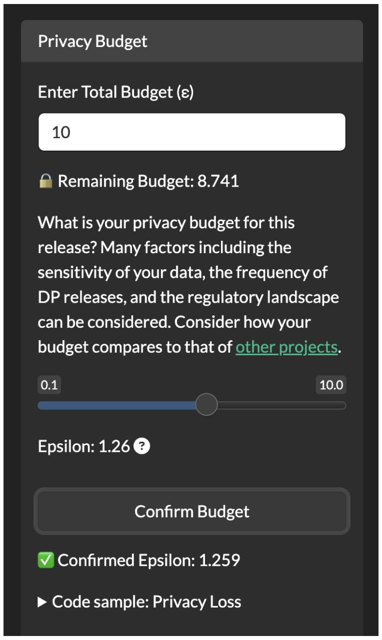
$$\Pr(D \mid y) = \frac{e^{\varepsilon}\, p_0}{e^{\varepsilon}\, p_0 + (1 - p_0)},$$

which we plot as the privacy risk (in percent), providing a tight, theory-grounde upper bound on how much an adversary's confidence can increase after observing the DP output.

Privacy Budget

DP Wizard currently prompts users to choose a privacy budget for a single query release using a slider. We extended this functionality by integrating a dynamic budget monitor into the interface where users can enter a total allowable privacy budget. When users select an ε value from 0.1-10 with the slider, the confirmation button must be clicked to finalize spending, reducing accidental loss. The confirmed ε value is then immediately deducted from the total allowable privacy budget. The remaining budget is displayed prominently, giving users continual feedback on their allocation decisions and helping them manage their cumulative privacy loss. This tracks the privacy cost of each query in real-time and alerts users when their budget is nearly depleted or fully exhausted by displaying a warning when ε exceeds the total allowable privacy budget. By having the interface track spending real-time, it encourages thoughtful allocation.

In future work, we would like to have each query (Histogram, Mean, Median, Count, Quantile and Standard Deviation) consume a specific portion of the budget based on the assigned ε. This supports more informed, iterative decision-making, particularly for analysts conducting multiple queries in an exploratory workflow.



Privacy Budget

Enter Total Budget (ε)

10

🔒 Remaining Budget: 8.741

What is your privacy budget for this release? Many factors including the sensitivity of your data, the frequency of DP releases, and the regulatory landscape can be considered. Consider how your budget compares to that of other projects.

0.1 10.0

Epsilon: 1.26 ?

Confirm Budget

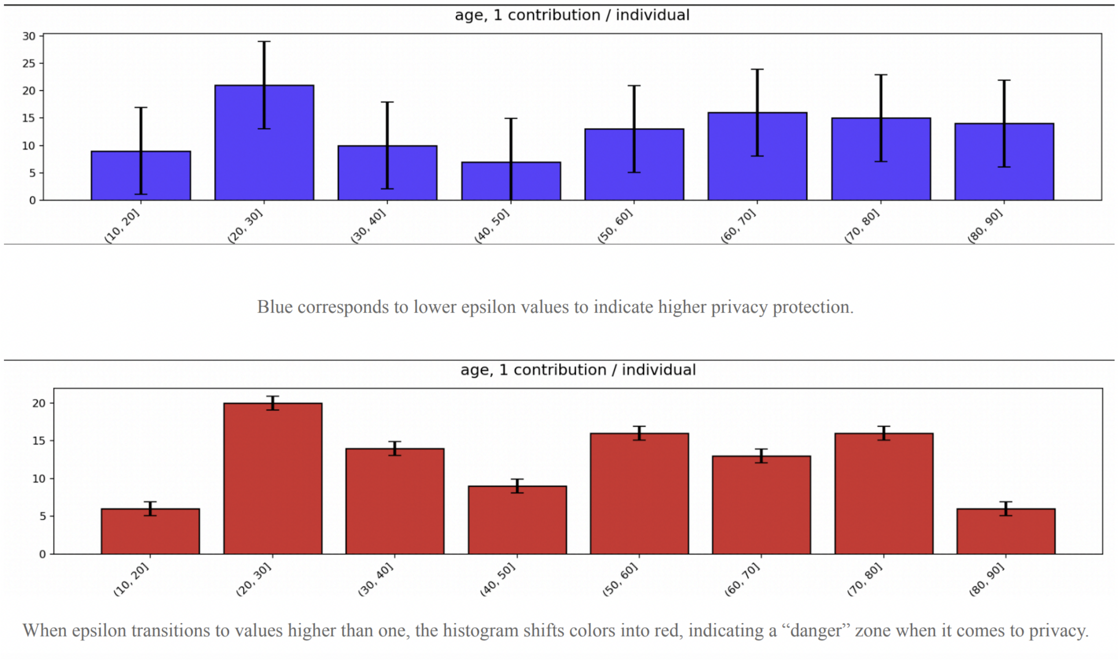
✅ Confirmed Epsilon: 1.259

▶ Code sample: Privacy Loss

Figure 8: Privacy Budget

User Interface Improvements

Recognizing that aesthetics and clarity are essential for reducing user intimidation and improving engagement, we made several improvements to the look and feel of the DP-Wizard interface. These aim to enhance the interpretability of differentially private outputs and foster a more comfortable learning experience for new users. We added a dark theme to the interface, not only to improve visual accessibility and contrast but also to make the site consistent with user interface conventions across data platforms. The theme also improved the appearance of buttons and widgets, and we added a clear page title named "DP Wizard Enhanced." To help users better comprehend the output distributions, we refined the histogram visualizations. Confidence intervals were enlarged and darkened to ensure visibility, and histogram bars were color-coded based on ε: for ε values under 1, we used shades of blue to signal higher privacy protection, while for ε values above 1, where privacy risk increases, the bars shift from light red to deep red, visually displaying increasing risk.



Blue corresponds to lower epsilon values to indicate higher privacy protection.

When epsilon transitions to values higher than one, the histogram shifts colors into red, indicating a “danger” zone when it comes to privacy.

Figure 9: Histograms Color-coded Based on Epsilon Value

This visual mapping gives users immediate cues about the privacy–utility tradeoff and reflects the risk modeling seen in Franzen et al. and DPP. When ε transitions to values higher than one, the histogram shifts into a red “danger” zone, signaling reduced privacy.

Results & Testing Usability

To evaluate the usability of our DP Wizard enhancements, we combined heuristic analysis with informal user testing, focusing on participants with and without technical backgrounds.

Heuristic Evaluation Using Nielsen’s 10 Usability Heuristics

Firstly, we examined our interface against Nielsen’s 10 Usability Heuristics—a widely recognized framework for user interface evaluation. Nielsen’s heuristics (1994) include principles such as visibility of system status, match between system and the real world, user control and freedom, and others that make it the ”most popular and frequently used set of usability heuristics that exist today,” according to The Decision Lab. Our evaluation results are as follows:

- **Visibility of System Status**
  - ϵ-slider instantly updates epsilon label, graph, and text
  - Budget warnings triggered immediately.
- **Match Between System & Real World**
  - User-friendly labels: “Privacy Budget,” “Unit of Privacy,” etc.

  - Familiar workflow to status quo for sheets navigation: CSV → Select columns → Set ϵ → Analyze.

- **User Control & Freedom**
  - Navigation buttons at top allow non-linear movement/flexibility.
  - Inputs (ϵ, columns) always editable or clearable.

- **Consistency & Standards**
  - Uniform card layout, inputs, and color conventions (blue = accuracy, red = risk).

- **Error Prevention**
  - Buttons disabled until valid inputs provided.
  - Warnings given for numeric bounds enforced and CSV mismatch.

- **Recognition Over Recall**
  - Current selections and ϵ values always visible so users do not have to remember prior selections.

- **Flexibility & Efficiency of Use**
  - Logarithmic ϵ-slider for quicker expert tuning.
  - Retains visual simplicity preserved for novice users.

- **Aesthetic & Minimalist Design**
  - Each card targets a specific function (e.g., Privacy Budget).
  - Elements are hidden if not needed at the time.

- **Help Recognize & Recover from Errors**
  - Feedback when budget is low/exhausted.
  - Inline messages to preclude attempts at navigation before sufficient criteria filled.

- **Help & Documentation**
  - "About" panel explains purpose and provides an issue-report link.
  - Hints and tooltips embedded throughout.

Basic User Study

After reviewing our program through these heuristics, we conducted an informal user study with four participants to tangibly test our new features. Our participants included two sophomore non-STEM roommates and two STEM-concentrating juniors without prior exposure to differential privacy. Altogether, we oriented our study to focus on testing our two design goals of: (1) empowering data analysts to perform realistic, privacy-preserving analysis, including a wider variety of DP query types and filters and iterative adjustment of parameters during analysis, and (2) enabling data analysts to understand what ϵ means in practice, visualizing the privacy-utility tradeoff, and tracking ϵ privacy budget usage during exploratory workflows. Participants completed the following tasks for both the original DP Wizard and our DP Wizard *Enhanced*.

Task 1: Data Analyst Scenario (Accessibility Focus)
Prompt:
Using the interface, report the mean and standard deviation of the 'grade' column, while managing your organization's privacy budget.

Follow-up Question:
On a scale from 1 to 10, how confidently could you use this tool for future statistical analyses requiring differential privacy?

Summary of Findings:

All of the participants rated the enhanced more highly than the original DP Wizard, with scores for the enhanced version ranging from 6 to 9. The participants without a technical background noted that the layout was "intuitive," the tooltips were "self-explanatory," and the analysis process was clearly walked-through. One user observed that "everything was either intuitive or well-explained in the hint bubbles," while another noted that the updated interface "made it easier to imagine using this in a real reporting job." The STEM participants highlighted the same features but nuanced that "larger-scale projects with more queries might make it harder to track how much privacy is being lost." Overall, Task 1 demonstrated that DP Wizard Enhanced further increased flexibility and lowered barriers to entry for more advanced workflows.

Task 2: Epsilon and Privacy Risk Exploration (Educational Focus)
Prompt:
Explore the privacy-utility graph and explanatory text. Adjust ϵ and describe what you learn about how privacy and accuracy trade off in differential privacy.

Follow-up Question:
In your own words, how would you now describe differential privacy and the role of epsilon?

Summary of Findings:

Across the board, participants cited the visualization and textual summary components as the most effective components for understanding how ϵ operates. One non-technical user said, "Before this, I had no idea what ϵ was. But now I know that if it's low, my results are fuzzier, and if it's high, they're sharper but riskier." Another remarked that the brief textual descriptions below the budget and visualization features provided "a real-world way to think about what you're giving up when you publish private stats." Likewise, STEM participants—who noted that they are typically accustomed to theoretical presentations—valued how the program bridged between both concrete numeric explanations as well of visual representations of how ϵ can improve adversary's guesses with higher values. One noted that they appreciated the clarification that the Accuracy and Privacy Risk values are simply "ballpark" estimates, highlighting how they anticipate this would be very helpful if attempting to manage risk as a novice DP user.

Summary

Across these tasks, our user feedback indicates that DP Wizard Enhanced improved analyst usability and educational clarity for users from varied technical backgrounds. The redesign made core DP con- cepts—particularly the privacy–utility tradeoff—concrete and approachable, directly supporting our goals of analyst accessibility and parameter comprehension. The most praised features were the interactive ϵ slider, dual-axis risk/accuracy visualization, and plain-language explanations of adversarial risk, which filled the gaps we noticed in the existing DP Communication software and literature. For example, only three existing studies have attempted to quantify Privacy Risk across varying epsilon values, and our interface drew from their mathematical methods to provide the first real-time-updated UI for demonstrating the Privacy-Utility Tradeoff and concurrent privacy budget implications. Moreover, we gleaned from our feedback that remaining challenges include clarifying downloadable outputs (for non-STEM-background users) and extending support for complex, repeated-query workflows.

References

Cummings, R., Kaptchuk, G., & Redmiles, E. M. (2021). "I need a better description": An investigation into user expectations for differential privacy. In *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security* (pp. 3037–3052). ACM.

Ding, B., Kulkarni, J., Yekhanin, S. (2021). Collecting Telemetry Data Privately. IEEE Transactions on Knowledge and Data Engineering, 33(1), 244–259. https://doi.org/10.1109/TKDE.2020.2970893

Franzen, D., Müller-Birn, C., & Wegwarth, O. (2024). Communicating the privacy-utility trade-off: Supporting informed data donation with privacy decision interfaces for differential privacy. *Proceedings of the ACM on Human-Computer Interaction, 8*(CSCW1), 1–56.

Gaboardi, M., Honaker, J., King, G., Murtagh, J., Nissim, K., Ullman, J., & Vadhan, S. (2016). PSI (): A private data sharing interface. *arXiv* preprint arXiv:1609.04340.

Nanayakkara, P., Kim, H., Wu, Y., Sarvghad, A., Mahyar, N., Miklau, G., & Hullman, J. (2024). Measure-observe-remeasure: An interactive paradigm for differentially-private exploratory analysis. *arXiv* preprint arXiv:2406.01964.

Nielsen, J., & Mack, R. L. (1994). *Usability inspection methods.* Wiley.

The Decision Lab. (n.d.). *Nielsen's heuristics.* Retrieved May 10, 2025, from `https://thedecisionlab.com/reference-guide/design/nielsens-heuristics`

Sarathy, V., Brown, J., Murtagh, J., Vadhan, S., & Hullman, J. (2023). DP Creator: Towards usable differential privacy for data analysis. *arXiv* preprint arXiv:2302.11775.

Zhang, D., Gaboardi, M., Hay, M., Miklau, G. (n.d.). A demonstration of the DPComp platform for empirical evaluation of differentially private algorithms. Retrieved from Zhang et al..

Member Contributions

- Yaying developed the functionality for the Count, Median and Standard Deviation queries, including backend logic and UI integration.

- Isaac implemented the epsilon tradeoff visualization graph and led the informal user evaluation based on Nielsen's usability heuristics.

- Brianna designed and implemented the privacy budget tracker, ensuring it reflected the remaining budget after the user confirms a set privacy budget for each query.

- Namat led the user interface redesign, including implementing a new color encoding for histograms to improve visual clarity and accessibility.

- Shiloh conducted the literature review, led the writing of the final report, and ensured all arguments were well-grounded in prior work and aligning with design goals.