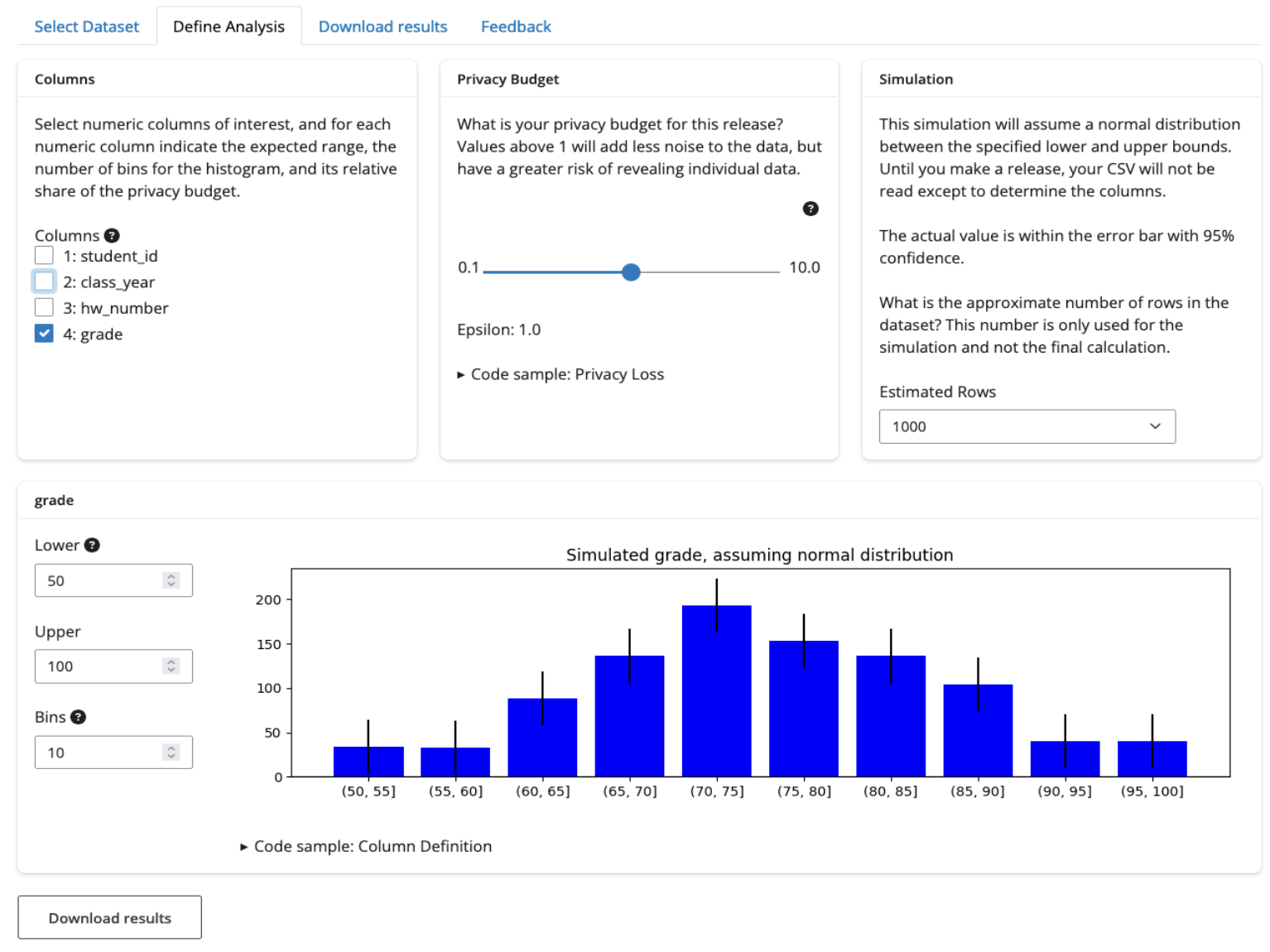
DP Wizard *Enhanced*

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

CS2080 Final Project

INTRODUCTION

Differential Privacy (DP) is a framework for computing statistics in a way that hides the presence or absence of any individual's data. It works by injecting noise into the released data. To help users explore and understand this concept, we utilize a DP Interface for interactive experimentation.



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

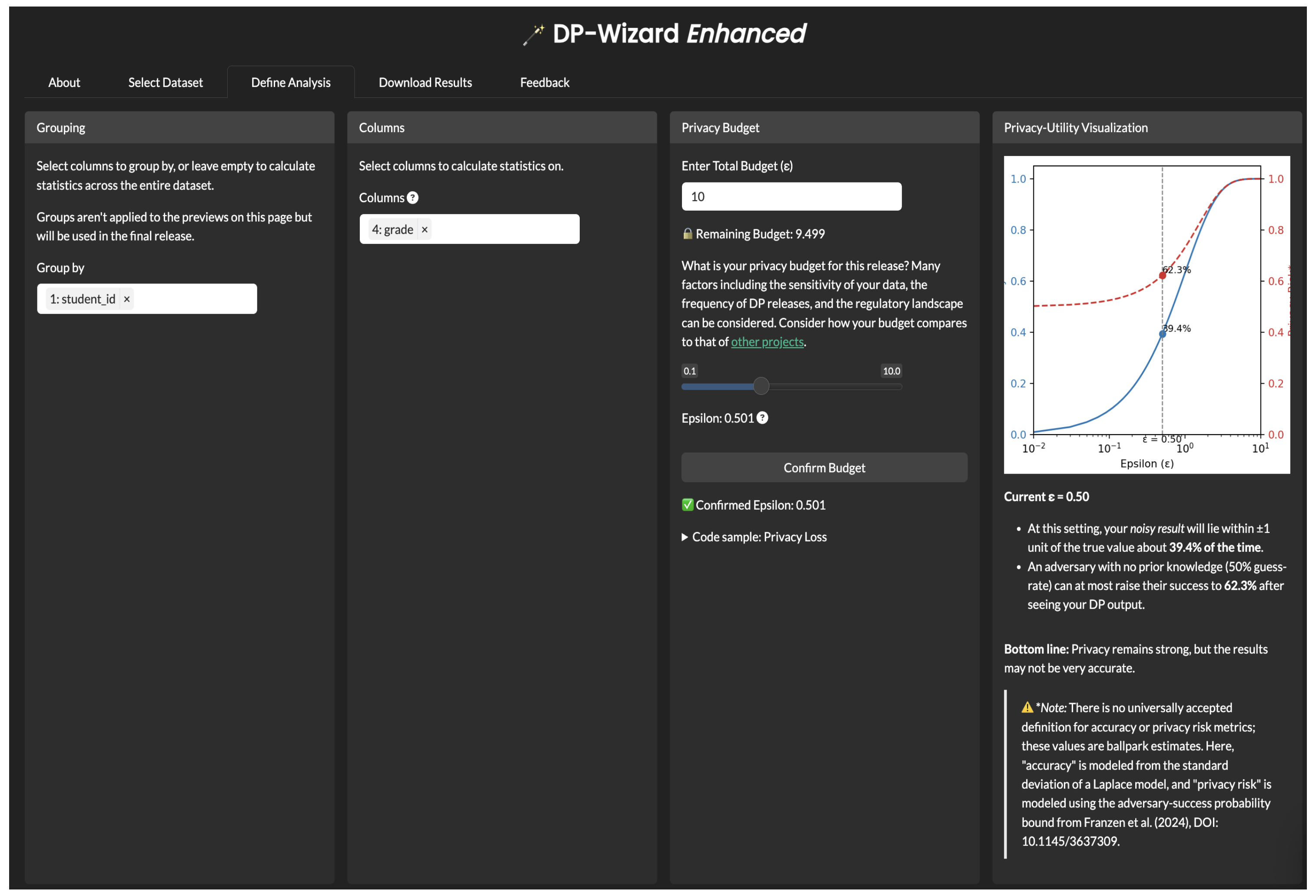
Bins: 10

Figure 1. Original DP Wizard Interface.

Design Objectives

Original DP Wizard design goals: (1) 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 builds on these foundations and is guided by two new design goals:

1. Enable data analysts to understand what ε means in practice, visualizing the privacy-utility tradeoff, and tracking ε privacy budget usage during exploratory workflows.
2. Empower 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.



Grouping

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 ×

Columns

Select columns to calculate statistics on.

Columns: 4: grade ×

Privacy Budget

Enter Total Budget (ε): 10

🔒 Remaining Budget: 9.499

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: 0.501

Confirm Budget

✅ Confirmed Epsilon: 0.501

▶ Code sample: Privacy Loss

Privacy-Utility Visualization

Current ε = 0.50

- At this setting, your *noisy result* will lie within ±1 unit of the true value about **39.4% of the time**.
- An adversary with no prior knowledge (50% guess-rate) can at most raise their success to **62.3%** after seeing your DP output.

Bottom line: Privacy remains strong, but the results may not be very accurate.

> ⚠ **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. DP-Wizard, Enhanced Interface.

BACKGROUND AND MOTIVATION

- Research highlights how misunderstood or overstated descriptions of differential privacy can lead to misplaced trust (Cummings et al., 2021; Smart et al., 2022).
- Users' trust depends not just on mathematical guarantees, but on clear communication about what DP protects, especially regarding control over data, transparency of use, and contextual relevance (Cummings et al., 2021).
- Usability studies (Dibia et al., 2024) reveal recurring challenges for DP tools, particularly around budgeting, statistical utility, and supporting users without privacy expertise.
- Tools like PSI (Gaboardi et al., 2016) and DP Creator (Sarathy et al., 2023) show that features like query diversity, privacy budget tracking, and intuitive explanations make DP interfaces more accessible and practical.
- Franzen et al. (2024) provide a model for visualizing privacy risk, showing how incremental changes to ε influence adversarial success rates.
- The importance of supporting exploratory workflows has been underscored by interfaces such as Overlook (Thaker et al., 2022) and the Measure-Observe-Remeasure paradigm (Nanayakkara et al., 2024), which allow incremental privacy spending and iterative data interaction.
- St. John et al. (2021) emphasize how intuitive outcome metrics (e.g., risk of guessing someone's data) can help non-experts make more informed privacy decisions.
- Thus, there is a clear need for enhanced DP tools that combine broader analytical functionality with communication strategies that align with user concerns, enabling data analysts to apply differential privacy confidently and meaningfully.

METHODOLOGY

Our enhanced interface supports 3 new query types: Count, Quantile, and Standard Deviation.

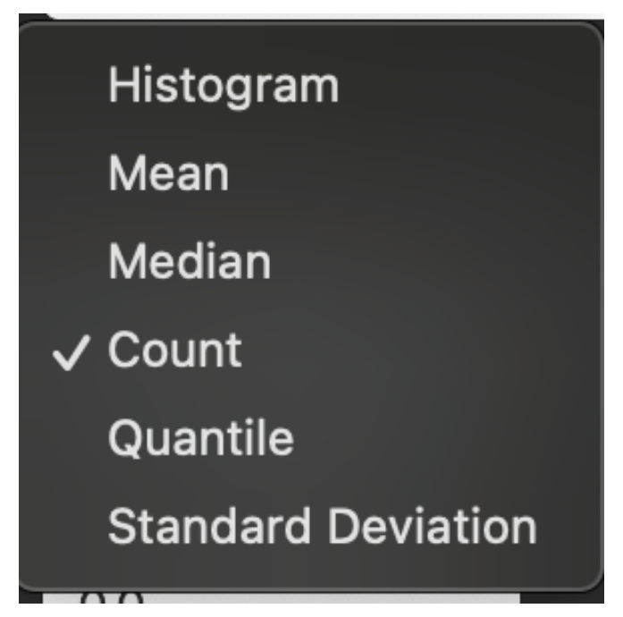
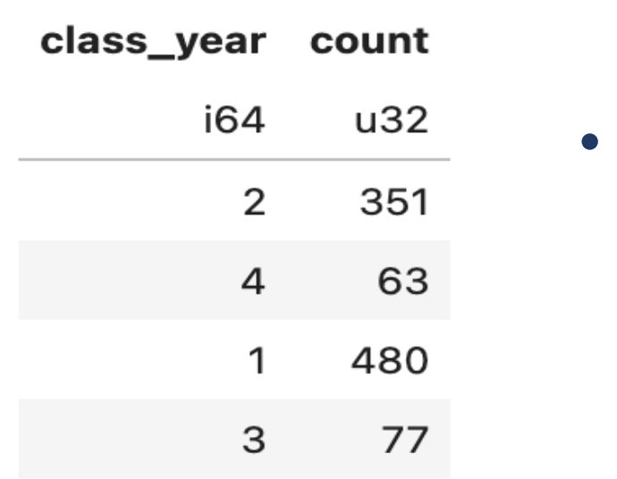


Figure 3. Dropdown Menu

This was accomplished by creating new folders for the 3 query types under the analysis directory. In particular, we implemented these 3 query types because of the benefits they provide data analysts.

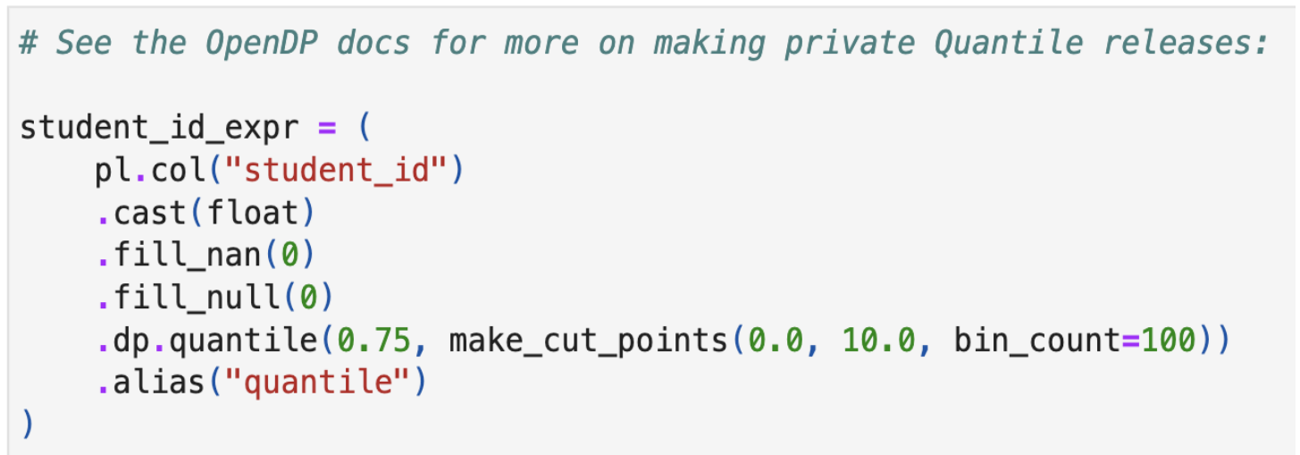
- **Count**: Count queries are simple yet foundational for initial data exploration and other query types. For example, knowing the count of the dataset gives us an immediate sense of scale of the dataset and helps us with filtering (ie. do we have enough observations in a subgroup to draw reliable conclusions?) 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).



class_year	count
i64	u32
2	351
4	63
1	480
3	77

Figure 4. Count Results

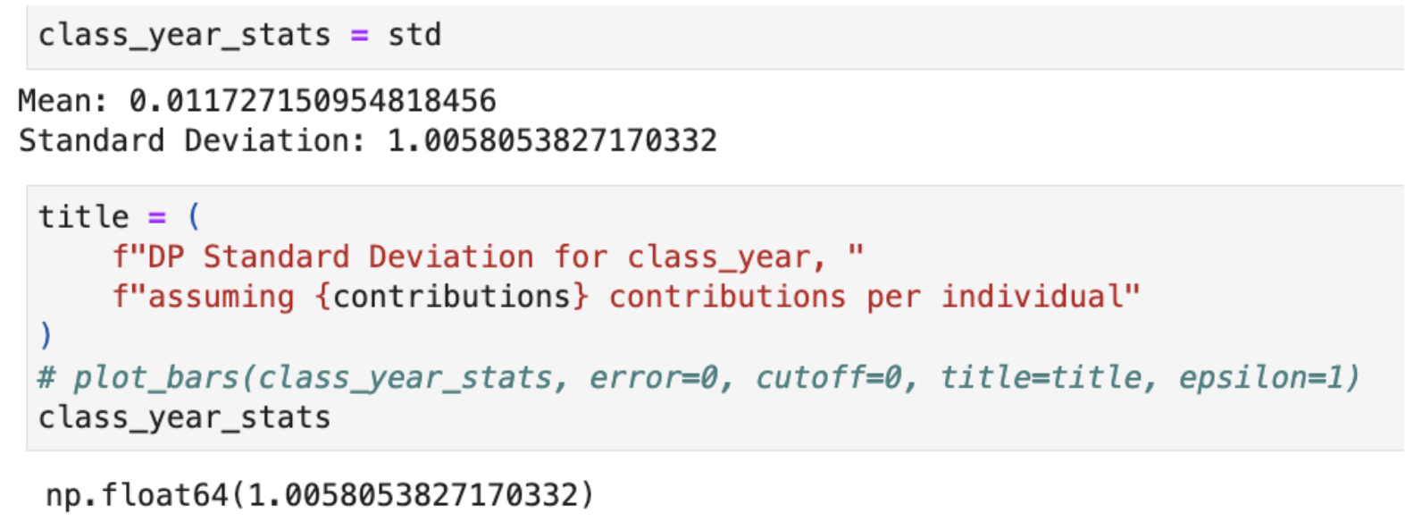
- **Quantile**: Quantile queries provide data analysts with distribution shape insights. The original DP-Wizard provides median (in other words, the 50th percentile) information, but our enhanced version provides information on 25th or 75th percentiles as well. In other words, users are more easily able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



```
# 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. Quantile Results



```
class_year_stats = std
```

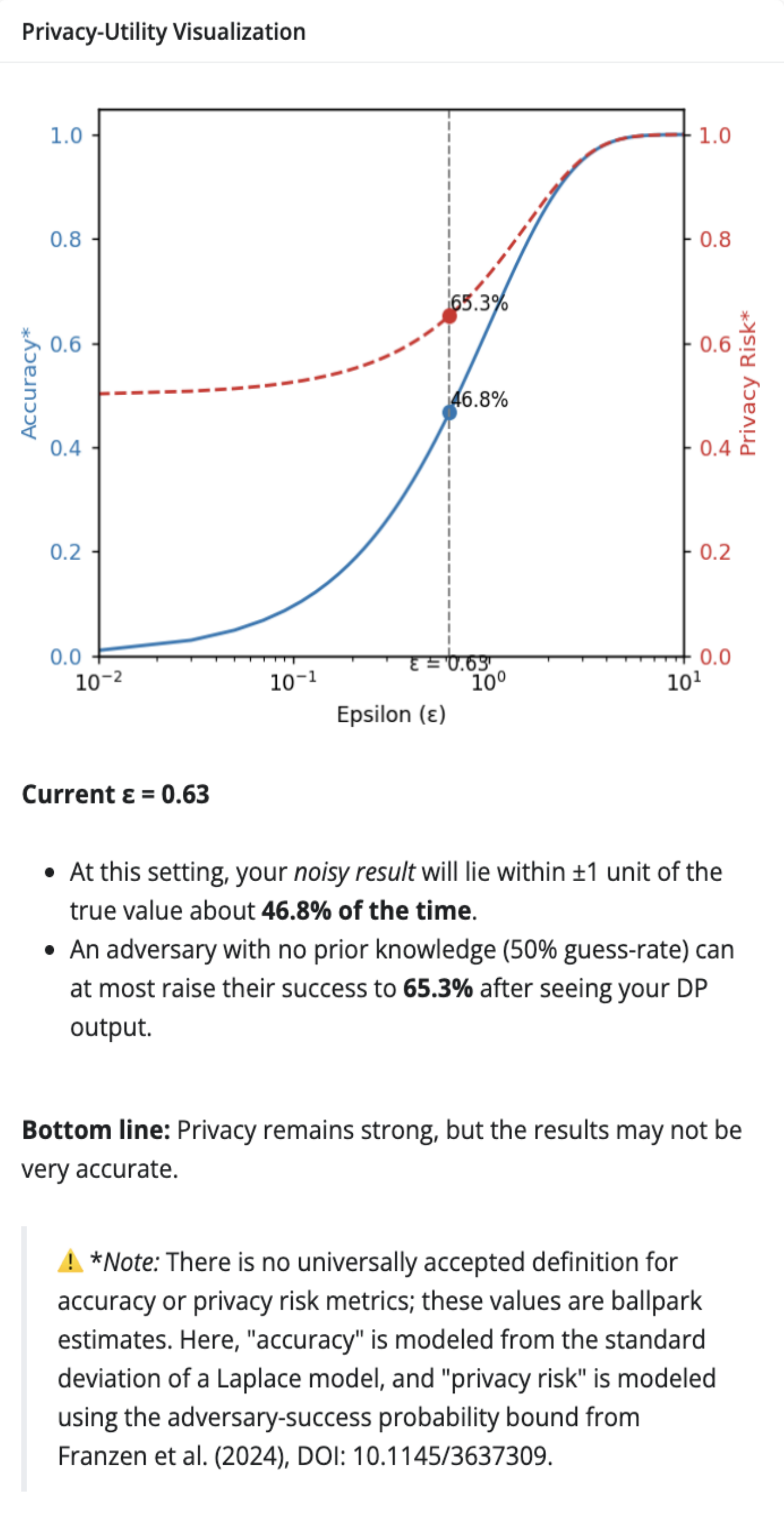
```
Mean: 0.011727150954818456
Standard Deviation: 1.0058053827170332
```

```
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. Standard Deviation Results

- **Standard Deviation**: Standard deviation queries are beneficial to data analysts because they quantify data points' variability from the mean. This helps analysts interpret fluctuations in the dataset and build private confidence intervals.



Privacy-Utility Visualization

Current ε = 0.63

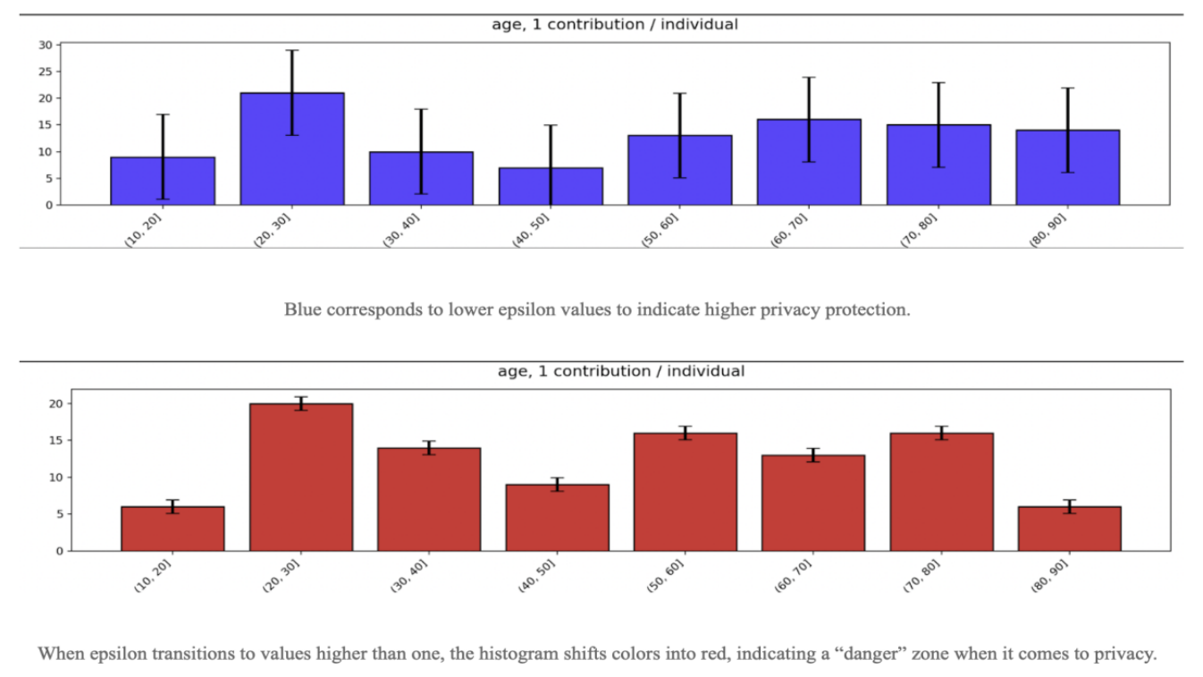
- At this setting, your *noisy result* will lie within ±1 unit of the true value about **46.8% of the time**.
- An adversary with no prior knowledge (50% guess-rate) can at most raise their success to **65.3%** after seeing your DP output.

Bottom line: Privacy remains strong, but the results may not be very accurate.

> ⚠ **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

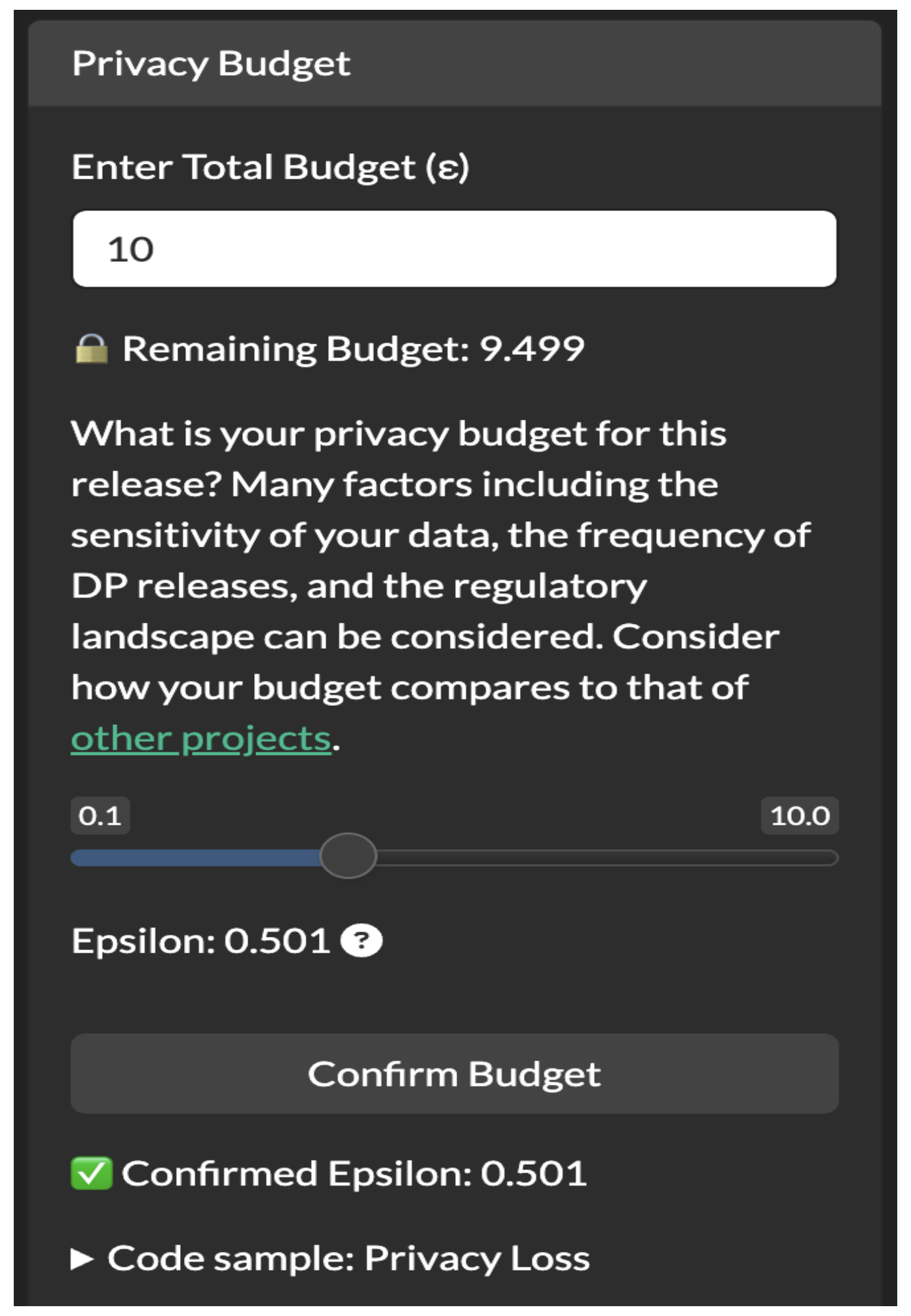
- **Epsilon Budget:** Users can enter a total allowable privacy budget, and the interface tracks spending across queries. When a user selects an ε value with the slider (0.1–10), a confirmation button must be clicked to finalize spending—reducing accidental loss. Real-time budget tracking and visual feedback encourage thoughtful allocation, warning users when their budget is low or exhausted.
- **Epsilon Visualization:** A dual-axis graph shows Accuracy (blue) and Privacy Risk (red) curves side by side with the ε slider. As users adjust ε, a vertical line moves and updates exact percentages at the chosen point, visually demonstrating the privacy-utility tradeoff. This design allows even non-experts to interpret DP impacts by simply aligning ε with a comfort zone on the chart.



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 8. Histograms, Original vs. New



Privacy Budget

Enter Total Budget (ε): 10

🔒 Remaining Budget: 9.499

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: 0.501

Confirm Budget

✅ Confirmed Epsilon: 0.501

▶ Code sample: Privacy Loss

Figure 9. Epsilon Privacy Budget

RESULTS & TESTING USABILITY

Heuristic Review (Nielsen's 10 Usability Heuristics)

We first ensured our interface conformed to core usability principles. DP Wizard Enhanced satisfied all ten of Nielsen's heuristics. Highlights include:

- **Visibility of System Status:** Epsilon label and graph update immediately with slider changes, budget warnings trigger live on screen
- **Match Between System and Real World:** User-friendly terms: "Privacy Budget," "Group by," "Unit of Privacy," intuitive flow: CSV → Columns → Parameters → Results
- **Help and Documentation:** Inline hints, tips, and "About" tab with clear description of tools/usage

User Study

We conducted a comparative usability study with **four participants**:

- Two roommate sophomores at Harvard concentrating in **non-STEM disciplines.**
- Two Harvard juniors in **STEM concentrations** without prior exposure to differential privacy.

Design Goal 1: Accessibility for Data Analysts

Task: *"Using the interface, report the **mean** and **standard deviation** of the 'grade' column, while managing your organization's privacy budget."*

Evaluation Prompt: *On a scale from 1 to 10, how confidently could you use this tool for future statistical analyses requiring differential privacy?*

Summary of Findings:

- Confidence ratings ranged from **6 to 9**.
- Users reported that the task flow was intuitive, with clear column selection and well-labeled inputs.
- The tooltip hints and immediate feedback (e.g., button enabling/disabling) helped users navigate comfortably.
- Concerns primarily focused on interpreting downloaded results and how well the interface would scale when navigating the privacy budget with multiple queries.

Design Goal 2: Communicating Privacy Tradeoffs and Epsilon

Task: *"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."*

Evaluation Prompt: *"In your own words, how would you now describe differential privacy and the role of epsilon?"*

Summary of Findings:

- Users found that the graph, especially when paired with plain-language text, clarified how ε affects both privacy risk and data utility.
- Participants were able to articulate that larger ε values result in more accurate outputs but less privacy, and vice versa.
- Recognized that ε is an abstract quantity, but felt the visual and numeric guidance made it approachable. STEM concentration respondents highlighted how "ballpark figures" for accuracy and privacy risk were still very helpful.

REFERENCES AND RELATED WORK

Rachel Cummings, Gabriel Kaptchuk, and Elissa M Redmiles. "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, pages 3037–3052, 2021.

Daniel Franzen, Claudia Müller-Birn, and Odette Wegwarth. Communicating the privacy-utility trade-off: Supporting informed data donation with privacy decision interfaces for differential privacy. Proceedings of the ACM on HumanComputer Interaction, 8(CSCW1):1–56, 2024.

Marco Gaboardi, James Honaker, Gary King, Jack Murtagh, Kobbi Nissim, Jonathan Ullman, and Salil Vadhan. Psi (ψ): A private data sharing interface. arXiv preprint arXiv:1609.04340, 2016.

Priyanka Nanayakkara, Hyeok Kim, Yifan Wu, Ali Sarvghad, Narges Mahyar, Gerome Miklau, and Jessica Hullman. Measure-observe-remeasure: An interactive paradigm for differentially-private exploratory analysis. arXiv preprint arXiv:2406.01964, 2024.

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.