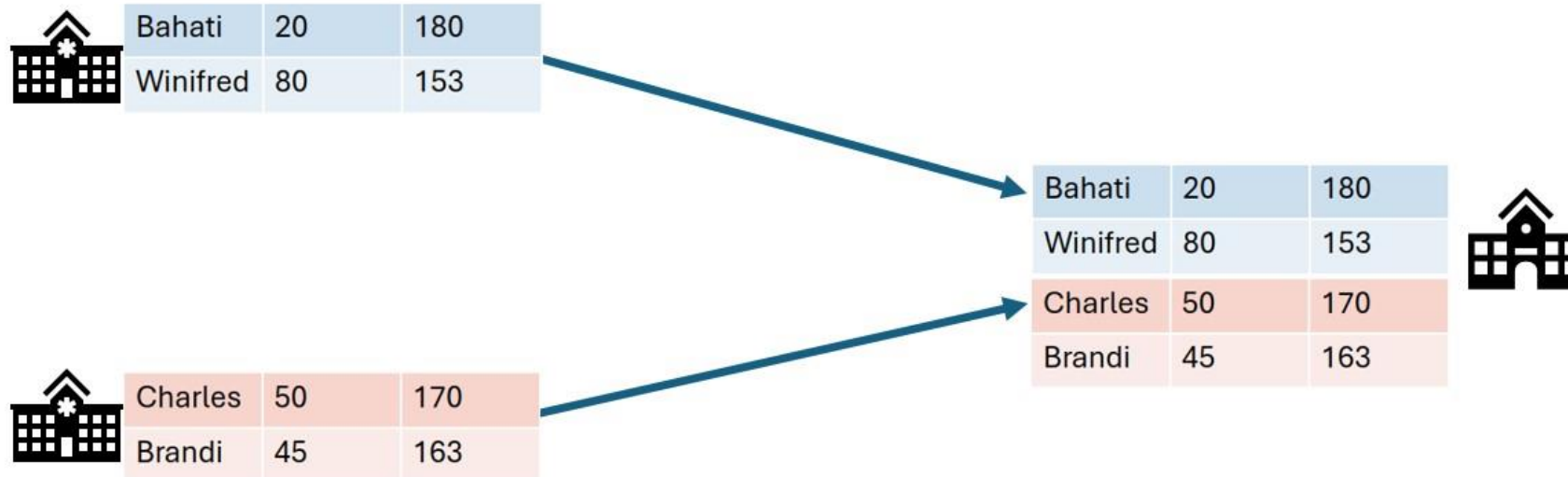
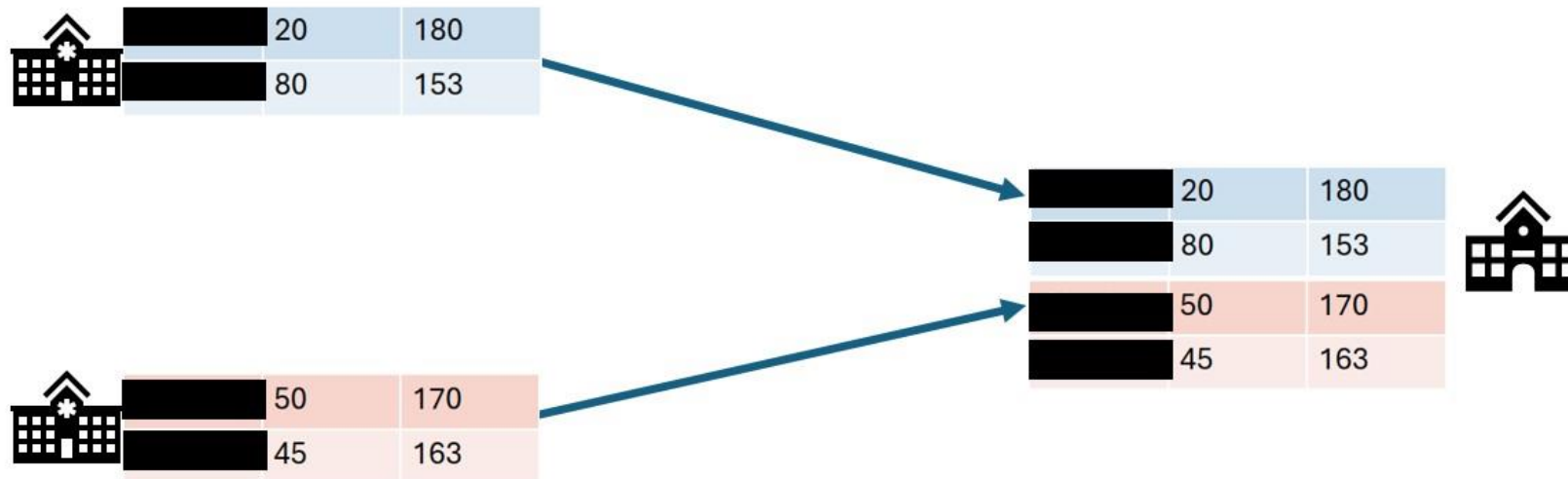


Introduction to Privacy Enhancing Technology

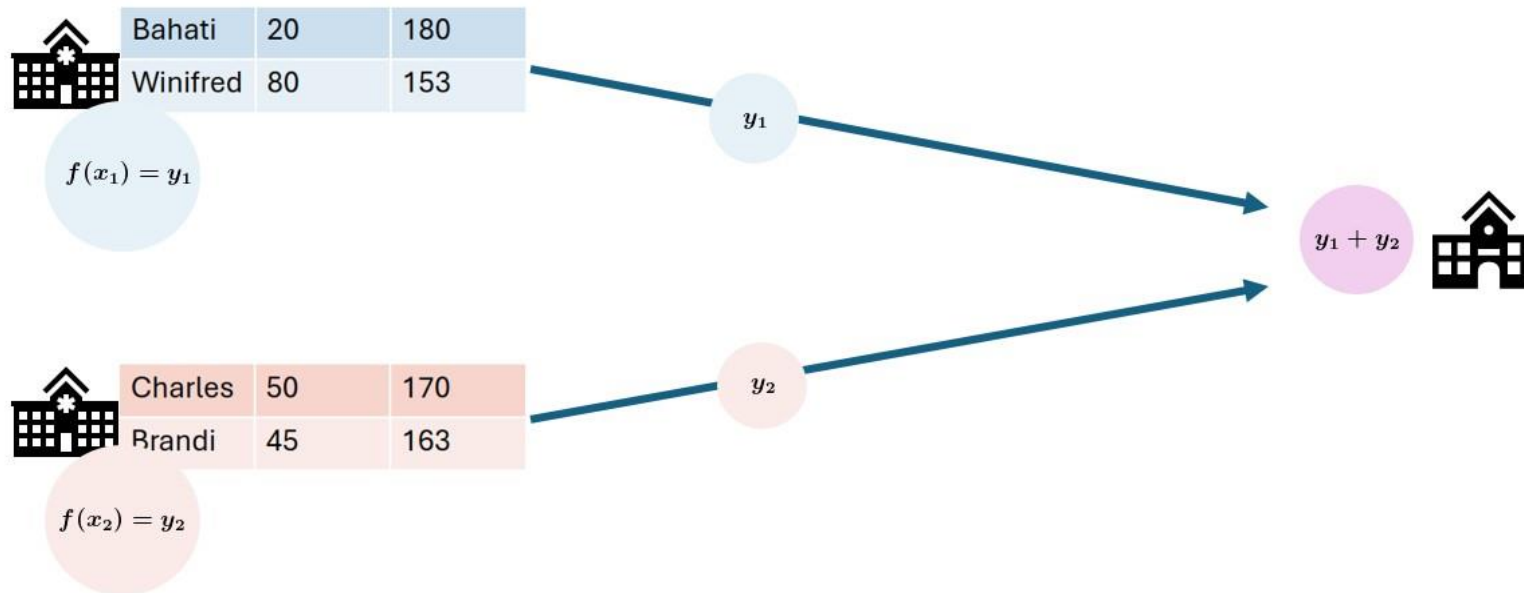
Classic analysis: Collect the data in one place



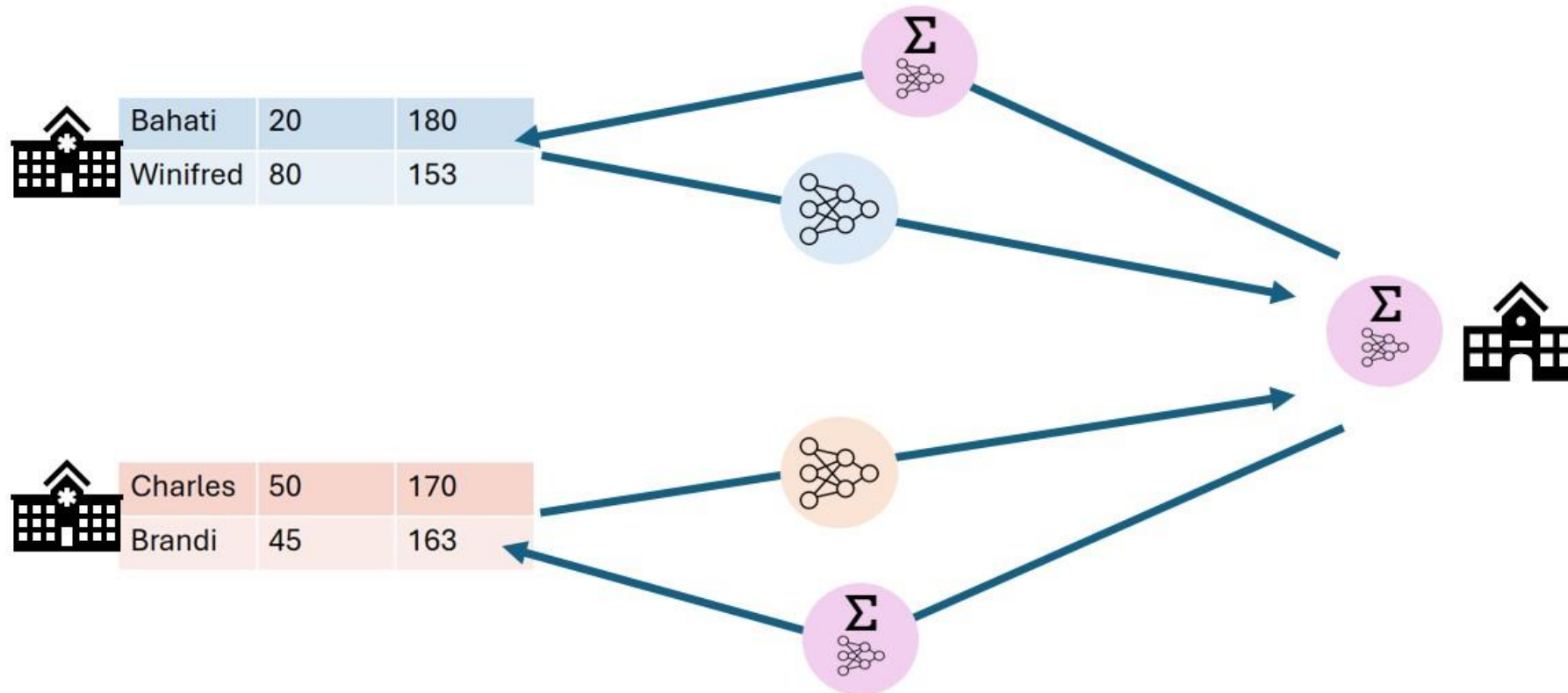
Data anonymization: Remove sensitive fields



Federated data analysis: Send analysis to the source, then aggregate

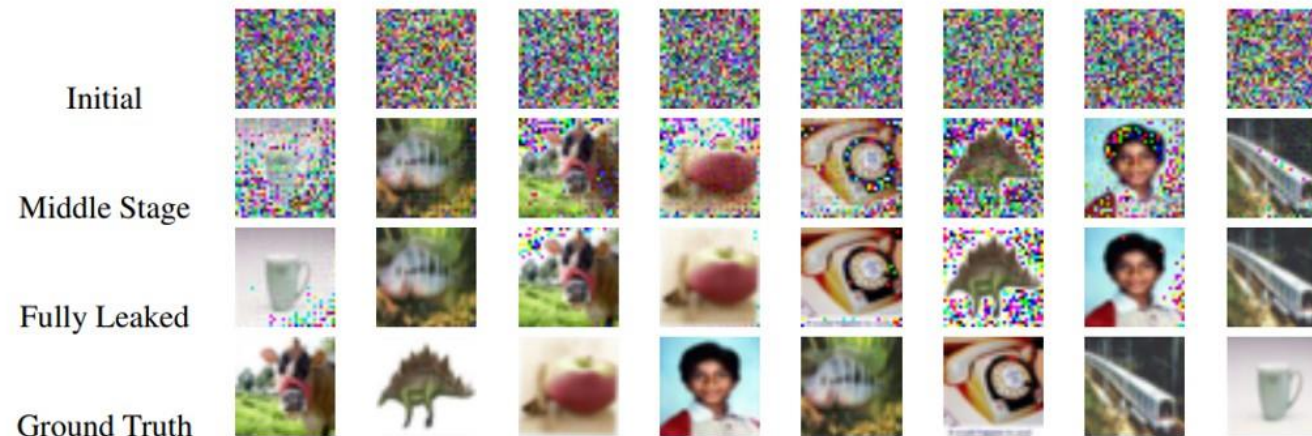


Federated learning: Train models locally, then aggregate



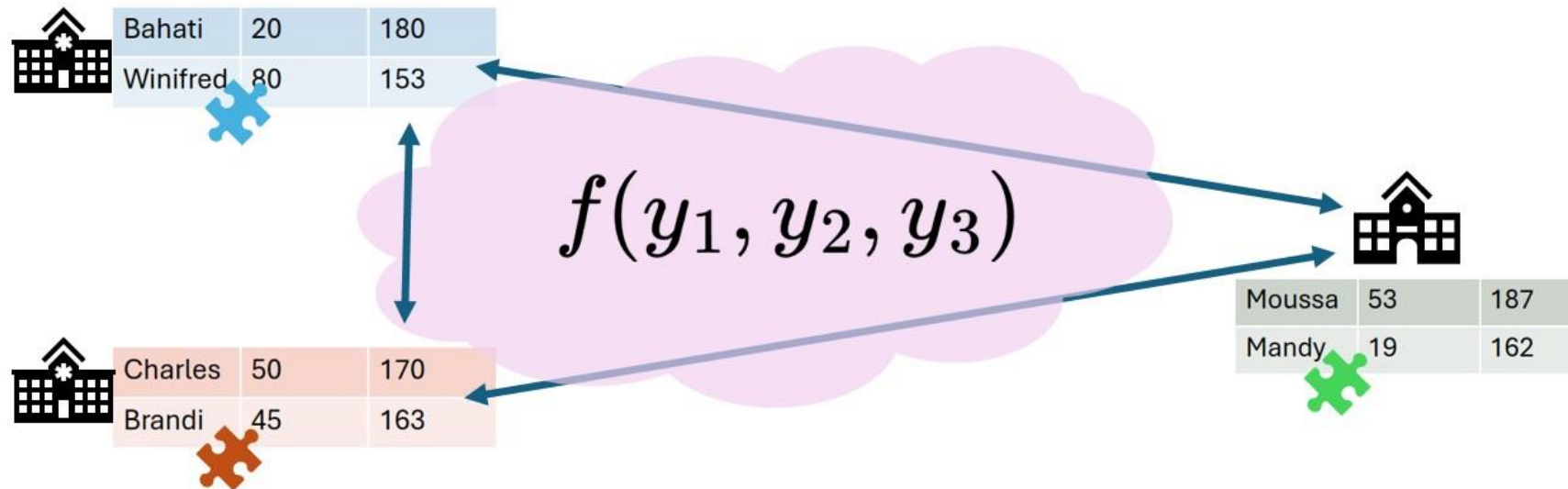
Federated learning and analysis can still leak data!

Gradient leakage



Zhu, Ligeng, Zhijian Liu, and Song Han. "Deep leakage from gradients."
Advances in neural information processing systems 32 (2019).

Secure multiparty computation: Send around encrypted puzzle pieces



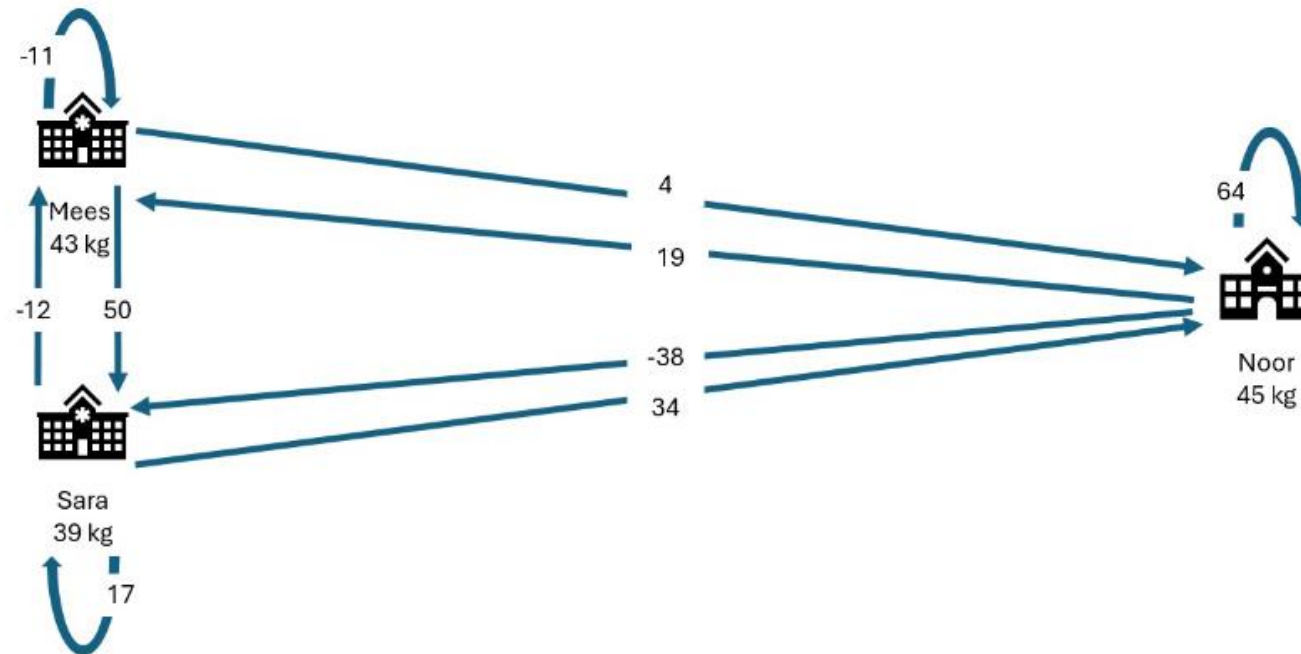
Secret sharing: an example



SECRET SHARING, AN EXAMPLE

Mees, Sara and Noor want to know how much they weigh in total. Mees weighs 43 kg, Sara weighs 39, Noor weighs 45. All three they think of 2 random numbers r_1 and r_2 so that $weight = r_1 + r_2 + x$. Finally they compute x by $x = weight - r_1 - r_2$

After computing the secret shares, they distribute these “cryptographical puzzle pieces” among their peers.



Secret sharing: an example



Mees

$$-11 - 12 + 19 = -4$$



Sara

$$50 + 17 - 38 = 29$$



Noor

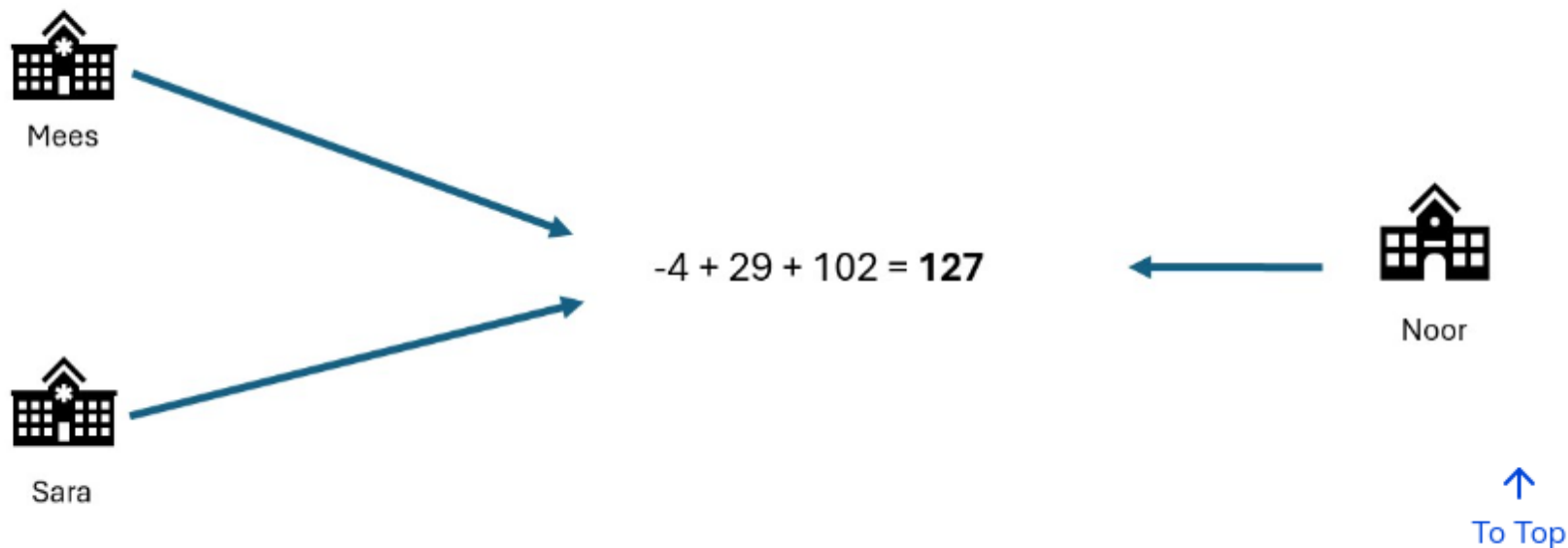
$$4 + 34 + 64 = 102$$



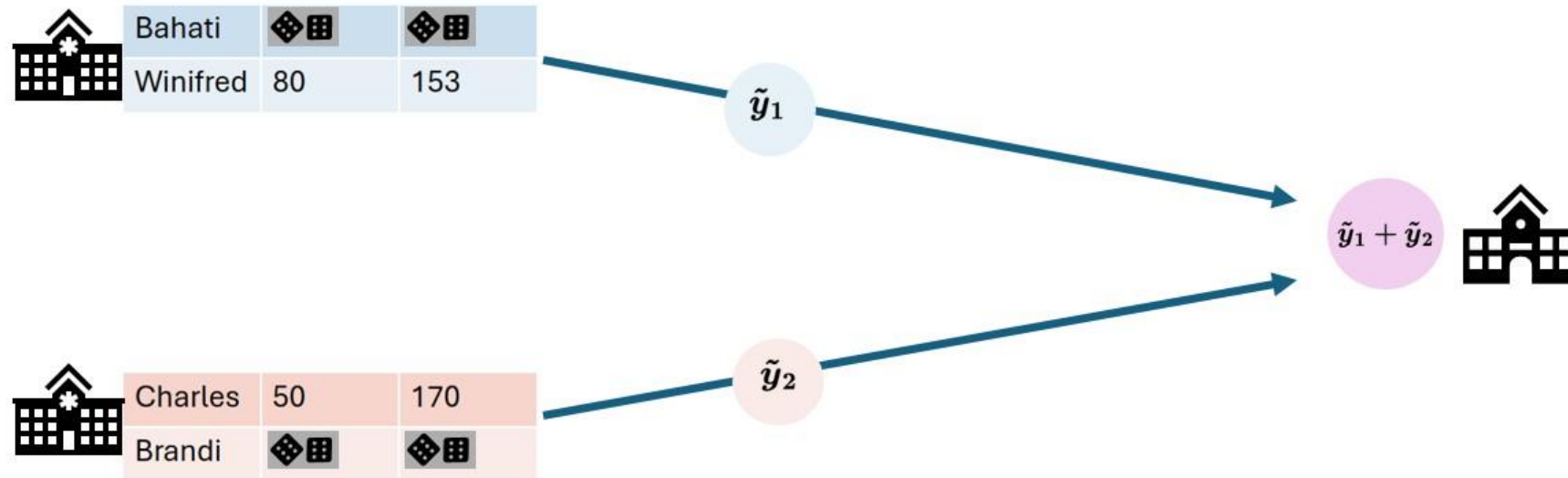
[To Top](#)

Secret sharing: an example

They add their sums together: $-4 + 29 + 102 = 127$ In this way, they have aggregated their data without sharing their individual data with anyone else.




Differential privacy: Add noise



Data partitioning

Horizontal partitioning



| | | |
|---------|----|-----|
| Charles | 50 | 170 |
| Brandi | 45 | 163 |
| Jez | 33 | 193 |
| Dionne | 25 | 168 |

Vertical partitioning

| | | | |
|---------|----|---------|-----|
| Charles | 50 | Charles | 170 |
| Brandi | 45 | Brandi | 163 |
| Jez | 33 | Jez | 193 |
| Dionne | 25 | Dionne | 168 |



Technology doesn't solve everything!

- Privacy enhancing technology is only a small part of the data sharing puzzle
- Some other factors
 - Trust
 - Regulations (either general or specific to the location)
 - Data harmonization

Summary

- With PETs you can derive insights from data without seeing individual records.
- PET analysis usually starts with the anonymization or pseudonymization of the data.
- In federated data analysis the analysis moves to the data, while in classic analysis the data moves around.
- In secure multiparty computation, computations are performed collaboratively without any one party being able to see all the raw data.
- Techniques from differential privacy add noise to the data to make it harder to reconstruct the original records from an aggregation.
- Privacy enhancing analyses usually stack multiple techniques on top of each other to provide multiple layers of protection.
- Horizontal partitioning means the records are split, while in vertical partitioning the features are split.
- In research on privacy sensitive data, technology is only one part of the story