Data Ethics and Privacy

PrepKC - 8th Grade

Who Am I?

Micah Melling

Chief Data Scientist at Americo Financial Life and Annuity

Masters Degree in Data Science from Rockhurst University

Professor at Park University

What are ways we can create <u>privacy and security</u> with data?

Privacy = Data can only be traced to a known individual by those allowed to do so

Security = Only allowed people can access the data

Masking

Masking

Replacing data to hide information

8161112222 → xxxxxxxxxx

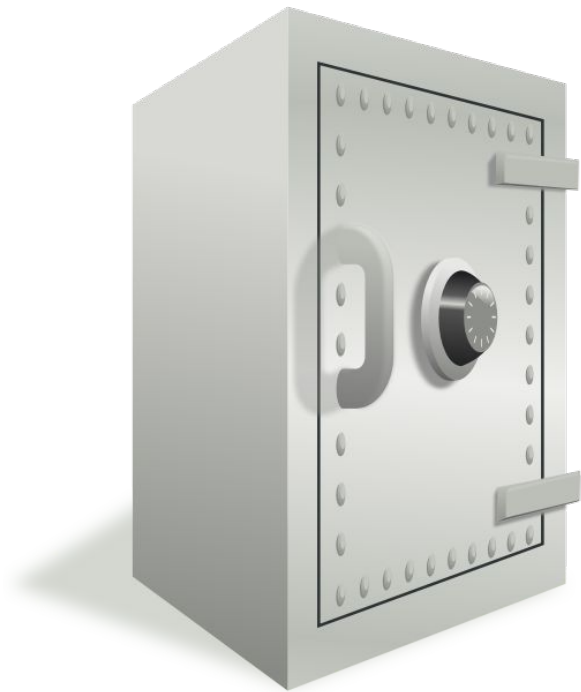Useful when you don't need the values for your work

Encryption

Encryption

Scrambling data that can be (potentially) unscrambled later

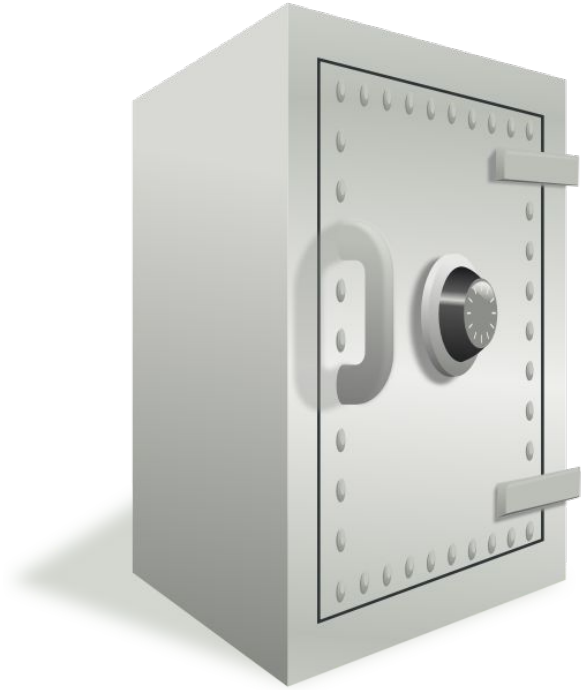https://emn178.github.io/online-tools/sha256.html

Useful when you need to retain uniqueness

Least-Privilege Access

Least-Privilege Access

Ensuring people only have access to data they need

What are ways we can act <u>ethically</u> with data?

Ethical = Doing what is right

We Want to Ensure Our Work is <u>Fair</u>

Disparate Impact

- Do adverse (negative) outcomes affect some groups more than others?

Types of Errors

- No analysis is perfect. (Does the average perfectly summarize every row of data?)
- Do the errors or assumptions of our analysis negatively impact some groups more than others?

We Need to Put the <u>Science</u> in Data Science

We need to perform rigorous analysis. What if sloppy work is used to impact people? Remember, we are performing science!

Sometimes you might feel pressure to come to a certain conclusion based on someone else's expectations. Let the data tell its story, even if it comes with a surprising or unwanted outcome.

Follow the Golden Rule

Demo Video: https://youtu.be/bqSOdQ5DQ8I