Session 06
Randomized Mechanism

General considerations

- In practice, there are strict rules and regulations what data can be collected, and how it can be used.
- Ref. GDPR
- Fines can be very high.
- Usually “I’ll just anonymize” is not good enough.
- Should you make it easy for big corporations to store your data?

Reminder

The randomized mechanism works as follows, on a **binary feature**, i.e. one that's distributed like $\mathrm{Bernoulli}(p)$, like e.g. gender.

1. Toss a coin $\propto \mathrm{Bernoulli}(\theta)$.
2. If the coin comes out heads, give the actual value (0 or 1).
3. Else, toss another (e.g. fair) coin, and give the result of that.

Reminder (cont.)

The resulting probability for returning 1 is

$$p' = \theta\, p + (1 - \theta)\frac{1}{2}$$

or to get an estimate for the original p,

$$p = \frac{1}{\theta}\left(p' - (1 - \theta)\frac{1}{2}\right)$$