> La privacy differenziale consente ai ricercatori e agli analisti di database di ottenere informazioni preziose dai database senza divulgare le informazioni di identificazione personale degli individui. Questo è fondamentale poiché molti database contengono una varietà di informazioni personali.

Il modo in cui funziona la privacy differenziale consiste nell'introdurre una **perdita di privacy o un parametro di budget per la privacy**, spesso indicato come epsilon (ε), nel set di dati. Questi parametri controllano quanto rumore o casualità viene aggiunto al set di dati non elaborato.

Ad esempio, si immagini di avere una colonna nel set di dati con le risposte "Sì"/"No" delle persone.

Si supponga ora di lanciare una moneta per ogni individuo:

- **Teste:** la risposta è lasciata così com'è.
- **Code:** si capovolge una seconda volta, registrando la risposta come "Sì" se testa e "No" se croce, indipendentemente dalla risposta reale.

Utilizzando questo processo, aggiungi casualità ai dati. Con una grande quantità di dati e le informazioni dal meccanismo di aggiunta del rumore, il set di dati rimarrà accurato in termini di misurazioni aggregate. La privacy entra in gioco consentendo a ogni singolo individuo di negare plausibilmente la propria vera risposta grazie al processo di randomizzazione.