Noise Injectors

Privacy Risk & Mitigation Summary

Identified Privacy Risks

- **Low k-Anonymity:** Unique Age + Gender + City combination in **22% of records**.

- **Poor l-Diversity:** Sensitive attribute *Disease* has only one value in **6 groups**.

- **High t-Closeness:** Income distribution deviates beyond the safe threshold in **4 groups**.

- **High δ-Disclosure Risk:** Re-identification probability exceeds **0.18** for rare occupations.

- **Background Knowledge Attack:** *Occupation + City* cross-reference reveals identities.

Applied Mitigation Measures

- **Generalized Age** into **5-year intervals** for broader anonymity.

- **Clustered Cities** into **regional categories** to hide small populations.

- **Added Controlled Noise** (±0.5%) to income values for differential privacy.

- **Pseudonymized Rare Occupations** to eliminate unique identifiers.

- **Generated Synthetic Data** for groups with elevated disclosure risk.

Utility & Compliance

- **Mean Absolute Error (Income):** 0.42% — high analytical accuracy maintained.

- **Legal Alignment:** Fully compliant with **DPDP Act 2023** privacy and utility standards.