Statistical Disclosure Control - Concepts

- Suda 2 is a recursive algorithm for finding **Minimal Sample Uniques**.
- The algorithm generates all possible variable subsets of defined categorical key variables and scans them for unique patterns in the subsets of variables.
- The lower the amount of variables needed to receive uniqueness, the higher the risk of the corresponding observation.

Statistical Disclosure Control - Concepts

- A new algorithm, SUDA2, is presented which finds minimally unique itemsets i.e., minimal itemsets of frequency one.
- These itemsets, referred to as Minimal Sample Uniques (MSUs), are important for statistical agencies who wish to estimate the risk of disclosure of their datasets.
- SUDA2 is a recursive algorithm which uses new observations about the properties of MSUs to prune and traverse the search space.
- Experimental comparisons with previous work demonstrate that SUDA2 is several orders of magnitude faster, enabling datasets of