

# Input specification

## 1. Input data

- Data
  - » Records: 200
  - » Attributes: 11

## 2. Attributes and transformations

- Insensitive attributes
  - » None
- Sensitive attributes
  - » race: String
- Identifying attributes
  - » idx: String
- Quasi-identifying attributes
  - » zipCode: String
  - » sex: String
  - » deathDate: String
  - » CBC:g/dL: String
  - » RC:%: String
  - » state: String
  - » diagDate: String
  - » birthDate: String
  - » age: String
- Response variables
  - » None
- Generalization hierarchy
  - » Attribute: zipCode
  - » Height: 3
  - » Minimum level: 0
  - » Maximum level: 2
- Generalization hierarchy
  - » Attribute: sex
  - » Height: 2
  - » Minimum level: 0
  - » Maximum level: 1
- Generalization hierarchy
  - » Attribute: deathDate
  - » Height: 3
  - » Minimum level: 0
  - » Maximum level: 2
- Generalization hierarchy
  - » Attribute: CBC:g/dL
  - » Height: 3
  - » Minimum level: 0
  - » Maximum level: 2
- Generalization hierarchy
  - » Attribute: RC:%
  - » Height: 3
  - » Minimum level: 0
  - » Maximum level: 2

- Generalization hierarchy
  - » Attribute: state
  - » Height: 3
  - » Minimum level: 0
  - » Maximum level: 2
- Generalization hierarchy
  - » Attribute: diagDate
  - » Height: 3
  - » Minimum level: 0
  - » Maximum level: 2
- Generalization hierarchy
  - » Attribute: birthDate
  - » Height: 3
  - » Minimum level: 0
  - » Maximum level: 2
- Generalization hierarchy
  - » Attribute: age
  - » Height: 5
  - » Minimum level: 0
  - » Maximum level: 4

### 3. Configuration

- Weights
  - » zipCode: 0.5
  - » race: 0.5
  - » sex: 0.5
  - » deathDate: 0.5
  - » CBC:g/dL: 1.0
  - » RC:%: 1.0
  - » state: 0.5
  - » idx: 0.5
  - » diagDate: 0.5
  - » birthDate: 0.5
  - » age: 0.5
- Settings
  - » Assume monotonicity: No
  - » Suppression limit: 1.0
- Risk thresholds
  - » Prosecutor risk: 0.09090909090909091
  - » Journalist risk: 0.09090909090909091
  - » Marketer risk: 0.09090909090909091

## Output properties

### 1. Output data

- Data
  - » Records: 200
  - » Attributes: 11
- Checksum
  - » SHA-256:  
1b16212526c07d93d0bf7366ef7a0a70d1549b2b5837ce2b7ea677ba014efcc0

## 2. Solutions

- Search space
  - » Size: 21870
  - » Materialized: 21870

## 3. Transformation

- Transformation
  - » Anonymity: Yes
  - » Minimum information loss: 5505.6744908194
  - » Maximum information loss: 5505.6744908194
  - Generalization scheme
    - » age: 3/4
    - » state: 2/2
    - » zipCode: 2/2
    - » sex: 1/1
    - » birthDate: 2/2
    - » diagDate: 2/2
    - » deathDate: 0/2
    - » CBC:g/dL: 2/2
    - » RC:%: 1/2

## 4. Data quality model

- Non-uniform entropy
  - » Aggregate function: Sum
  - » Monotonic: No
  - » Generalization factor: 1.0
  - » Suppression factor: 1.0

## 5. Privacy models

- t-Closeness
  - » Attribute: race
  - » Threshold (t): 0.5
  - » Distance: Hierarchical
- k-Anonymity
  - » Threshold (k): 11