

Project

1. Properties

- Project
 - » Name: Adult Salary
 - » Description: Not specified

Input specification

1. Input data

- Data
 - » Records: 48842
 - » Attributes: 14

2. Attributes and transformations

- Insensitive attributes
 - » edunum: Integer
 - » capitalloss: *Integer*
 - » nativecountry: *String*
 - » education: String
 - » occupation: String
 - » hoursperworks: Integer
 - » relationship: String
 - » capitalgain: *Integer*
- Sensitive attributes
 - » salaryclass: *String*
- Identifying attributes
 - » None
- Quasi-identifying attributes
 - » maritalstatus: *String*
 - » workclass: String
 - » race: String
 - » sex: String
 - » age: Integer
- Response variables
 - » None
- Generalization hierarchy
 - » Attribute: race
 - » Height: 4
 - » Minimum level: 0
 - » Maximum level: 3
- Generalization hierarchy
 - » Attribute: sex
 - » Height: 2
 - » Minimum level: 0
 - » Maximum level: 1
- Generalization hierarchy
 - » Attribute: maritalstatus
 - » Height: 3
 - » Minimum level: 0

- » Maximum level: 2
- Generalization hierarchy
 - » Attribute: workclass
 - » Height: 3
 - » Minimum level: 0
 - » Maximum level: 2
- Generalization hierarchy
 - » Attribute: age
 - » Height: 4
 - » Minimum level: 0
 - » Maximum level: 3

3. Configuration

- Weights
 - » maritalstatus: 0.5
 - » education: 0.5
 - » workclass: 0.5
 - » race: 0.5
 - » sex: 0.5
 - » age: 0.5
- Settings
 - » Assume monotonicity: No
 - » Suppression limit: 0.0
- Risk thresholds
 - » Prosecutor risk: 0.1
 - » Journalist risk: 0.1
 - » Marketer risk: 0.1

Output properties

1. Output data

- Data
 - » Records: 48842
 - » Attributes: 14
- Checksum
 - » SHA-256:
0bd6ffef12189b2927216ef1c0ab35bd82718fe16ff42bf7681798d6f93844e1

2. Solutions

- Search space
 - » Size: 288
 - » Materialized: 64

3. Transformation

- Transformation
 - » Anonymity: Yes
 - » Minimum information loss: 0.2744334291740087
 - » Maximum information loss: 0.2744334291740087
 - Generalization scheme
 - » age: 2/3
 - » workclass: 1/2

- » maritalstatus: 2/2
- » race: 2/3
- » sex: 0/1

4. Data quality model

- Loss
 - » Aggregate function: Geometric mean
 - » Monotonic: Yes
 - » Generalization factor: 1.0
 - » Suppression factor: 1.0

5. Privacy models

- k-Anonymity
 - » Threshold (k): 10
- t-Closeness
 - » Attribute: salaryclass
 - » Threshold (t): 0.4999995009999999
 - » Distance: Equal

ARX