

# 数据隐私实验报告

魏钊 PB18111699

## 一 . Samarati 算法

### 1 . 问题描述

使用 Samarati (基于二分法) 算法完成 k 匿名。

要求:

$QI = \{\text{age, gender, race, marital\_status}\}$  (categorical 型),  $S = \{\text{occupation}\}$

输入: data, k, maxSup (data 是数据集, k 是 K-Anonymity 的参数, maxSup 表示最大 suppression 的个数);

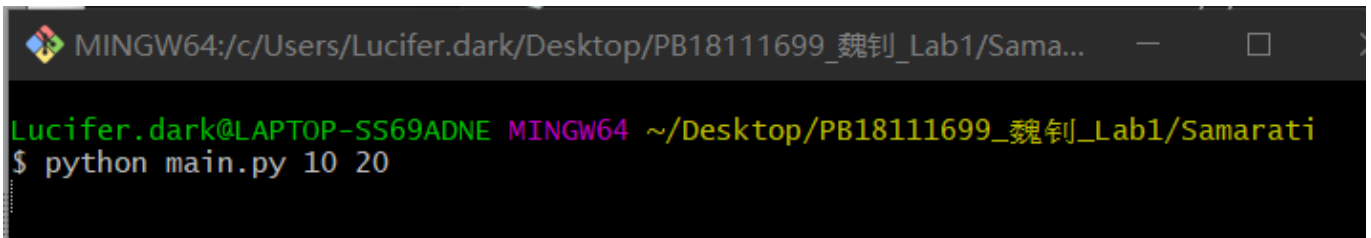
输出: 匿名后的数据集。

评价指标: 运行时间和 LM

### 2 . 程序使用指南以及程序分析

在 Samarati 文件夹下在终端执行 main.py 文件。

例如:  $K=10, \text{Maxsup}=20$

A terminal window with a dark background. The title bar shows the path 'MINGW64:/c/Users/Lucifer.dark/Desktop/PB18111699\_魏钊\_Lab1/Sama...'. The prompt is 'Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699\_魏钊\_Lab1/Samarati'. The command entered is '\$ python main.py 10 20'.

注意: 程序调用了 pandas

程序的输出为 Samarati 文件下的 Final.txt(泛化后的数据), 和终端中打印的三行数据。

三行数据分别为:

划分向量[age,Gender,race, Marital\_status] 注: 0 为不泛化输出原始数据

运行时间

例如：



The screenshot shows a Windows command prompt window with the title bar "MINGW64:/c/Users/Lucifer.dark/Desktop/PB18111699\_魏钊\_Lab1/Sama...". The prompt is "Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699\_魏钊\_Lab1/Samarati". The user has entered the command "\$ python main.py 10 20". The output is "[2, 1, 0, 1]", "LM=1.226823", and "Running time: 16.22192621231079 Seconds".

```
MINGW64:/c/Users/Lucifer.dark/Desktop/PB18111699_魏钊_Lab1/Sama...  
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Samarati  
$ python main.py 10 20  
[2, 1, 0, 1]  
LM=1.226823  
Running time: 16.22192621231079 Seconds
```

相关的程序分析在代码注释中进行了详细的叙述。

### 3. 实验结果分析

$K=1, \maxSup=0$  时, 数据不泛化, 即为原始数据。

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Samarati
$ python main.py 1 0
[0, 0, 0, 0]
LM=0.000000
Running time: 11.600184202194214 Seconds
```

Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

77,Male,Other,Married-civ-spouse,Machine-op-inspct  
 80,Female,Amer-Indian-Eskimo,Widowed,Other-service  
 71,Male,Amer-Indian-Eskimo,Married-civ-spouse,Farming-fishing  
 90,Female,Black,Divorced,Sales  
 90,Male,Black,Never-married,Other-service  
 90,Male,Black,Never-married,Exec-managerial  
 81,Female,Black,Widowed,Priv-house-serv  
 84,Male,Black,Married-civ-spouse,Prof-specialty  
 78,Female,Black,Widowed,Other-service  
 76,Female,Black,Widowed,Sales  
 77,Female,Black,Married-civ-spouse,Priv-house-serv  
 71,Female,Black,Divorced,Priv-house-serv  
 71,Male,Black,Separated,Sales  
 71,Male,Black,Widowed,Craft-repair  
 71,Male,Black,Married-civ-spouse,Craft-repair  
 74,Female,Black,Widowed,Adm-clerical  
 74,Female,Black,Widowed,Other-service  
 74,Female,Black,Widowed,Priv-house-serv

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

K=10,maxsup=0 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Samarati
$ python main.py 10 0
[2, 0, 0, 2]
LM=1.113924
Running time: 15.087336778640747 Seconds
```

Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

81~90,Female,White,\*,Other-service  
81~90,Female,White,\*,Adm-clerical  
81~90,Female,White,\*,Adm-clerical  
81~90,Female,White,\*,Other-service  
81~90,Female,White,\*,Exec-managerial  
81~90,Female,White,\*,Prof-specialty  
81~90,Female,White,\*,Exec-managerial  
81~90,Female,Black,\*,Sales  
81~90,Female,White,\*,Prof-specialty  
81~90,Female,White,\*,Adm-clerical  
81~90,Female,White,\*,Adm-clerical  
81~90,Female,White,\*,Adm-clerical  
81~90,Male,Black,\*,Other-service  
81~90,Male,White,\*,Prof-specialty  
81~90,Male,White,\*,Exec-managerial  
81~90,Male,Asian-Pac-Islander,\*,Other-service  
81~90,Male,White,\*,Handlers-cleaners  
81~90,Male,Asian-Pac-Islander,\*,Other-service

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

K=10, maxSup=20 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Samarati
$ python main.py 10 20
[2, 1, 0, 1]
LM=1.226823
Running time: 15.392743825912476 Seconds
```

Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

81~90,\*,Black,NM,Other-service  
81~90,\*,White,NM,Other-service  
81~90,\*,Asian-Pac-Islander,NM,Other-service  
81~90,\*,White,NM,Handlers-cleaners  
81~90,\*,Asian-Pac-Islander,NM,Other-service  
81~90,\*,White,NM,Adm-clerical  
81~90,\*,Black,NM,Exec-managerial  
81~90,\*,White,NM,Other-service  
81~90,\*,White,NM,Exec-managerial  
81~90,\*,White,NM,Prof-specialty  
81~90,\*,White,NM,Exec-managerial  
81~90,\*,Asian-Pac-Islander,NM,Prof-specialty  
81~90,\*,White,NM,Adm-clerical  
81~90,\*,White,NM,Exec-managerial  
71~80,\*,White,NM,Machine-op-inspct  
71~80,\*,White,NM,Prof-specialty  
71~80,\*,White,NM,Exec-managerial  
71~80,\*,White,NM,Craft-repair

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

K=100, maxsup=50 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏判_Lab1/Samarati
$ python main.py 100 50
[3, 0, 1, 0]
LM=1.240506
Running time: 11.940584182739258 Seconds
```

Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

71~90,Female,\*,Widowed,Adm-clerical  
71~90,Female,\*,Divorced,Adm-clerical  
71~90,Female,\*,Widowed,Adm-clerical  
71~90,Female,\*,Widowed,Other-service  
71~90,Female,\*,Never-married,Other-service  
71~90,Female,\*,Widowed,Other-service  
71~90,Female,\*,Divorced,Tech-support  
71~90,Female,\*,Married-civ-spouse,Adm-clerical  
71~90,Female,\*,Widowed,Sales  
71~90,Female,\*,Married-spouse-absent,Prof-specialty  
71~90,Female,\*,Married-civ-spouse,Exec-managerial  
71~90,Female,\*,Widowed,Priv-house-serv  
71~90,Female,\*,Widowed,Transport-moving  
71~90,Female,\*,Widowed,Adm-clerical  
71~90,Female,\*,Widowed,Adm-clerical  
71~90,Female,\*,Separated,Adm-clerical  
71~90,Female,\*,Widowed,Priv-house-serv  
71~90,Female,\*,Widowed,Adm-clerical

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

K=100, maxsup=200 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏判_Lab1/Samarati
$ python main.py 100 200
[3, 0, 1, 0]
LM=1.240506
Running time: 14.734896183013916 Seconds
```

Final - 记事本

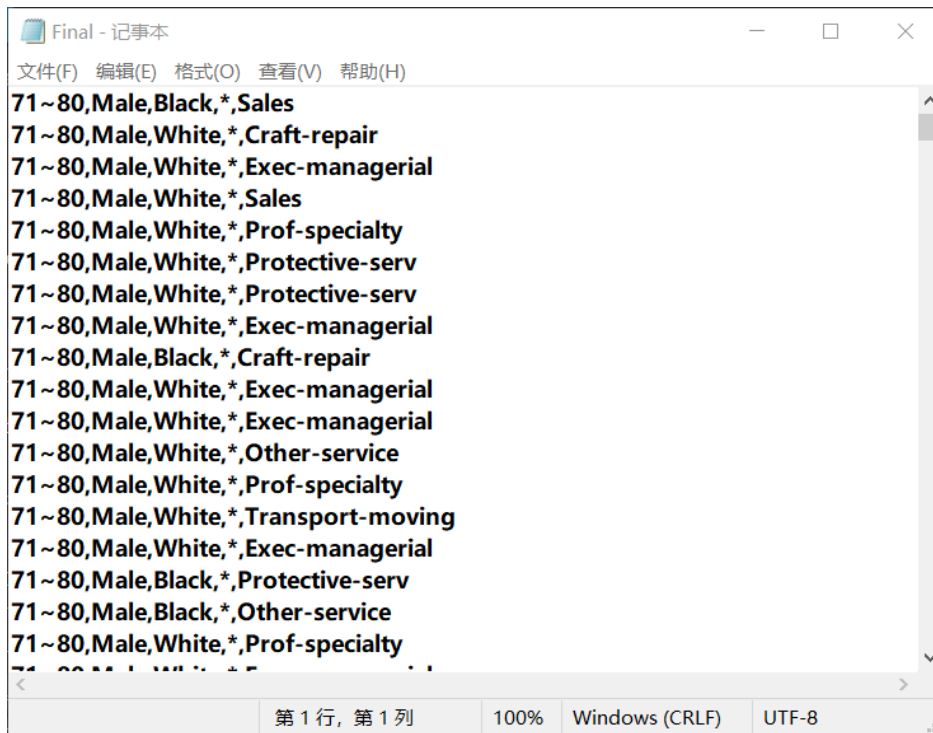
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

71~90,Male,\*,Married-civ-spouse,Prof-specialty  
71~90,Male,\*,Married-civ-spouse,Exec-managerial  
71~90,Male,\*,Separated,Sales  
71~90,Female,\*,Widowed,Adm-clerical  
71~90,Male,\*,Married-civ-spouse,Sales  
71~90,Male,\*,Married-civ-spouse,Craft-repair  
71~90,Female,\*,Divorced,Adm-clerical  
71~90,Female,\*,Widowed,Adm-clerical  
71~90,Male,\*,Married-civ-spouse,Craft-repair  
71~90,Male,\*,Married-civ-spouse,Exec-managerial  
71~90,Male,\*,Married-civ-spouse,Sales  
71~90,Female,\*,Widowed,Other-service  
71~90,Male,\*,Married-civ-spouse,Prof-specialty  
71~90,Female,\*,Widowed,Other-service  
71~90,Female,\*,Divorced,Tech-support  
71~90,Female,\*,Married-civ-spouse,Adm-clerical  
71~90,Female,\*,Widowed,Sales  
71~90,Male,\*,Married-civ-spouse,Protective-serv

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

K=100, maxsup=500 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Samarati
$ python main.py 100 500
[2, 0, 0, 2]
LM=1.113924
Running time: 15.541531801223755 Seconds
```



Final - 记事本

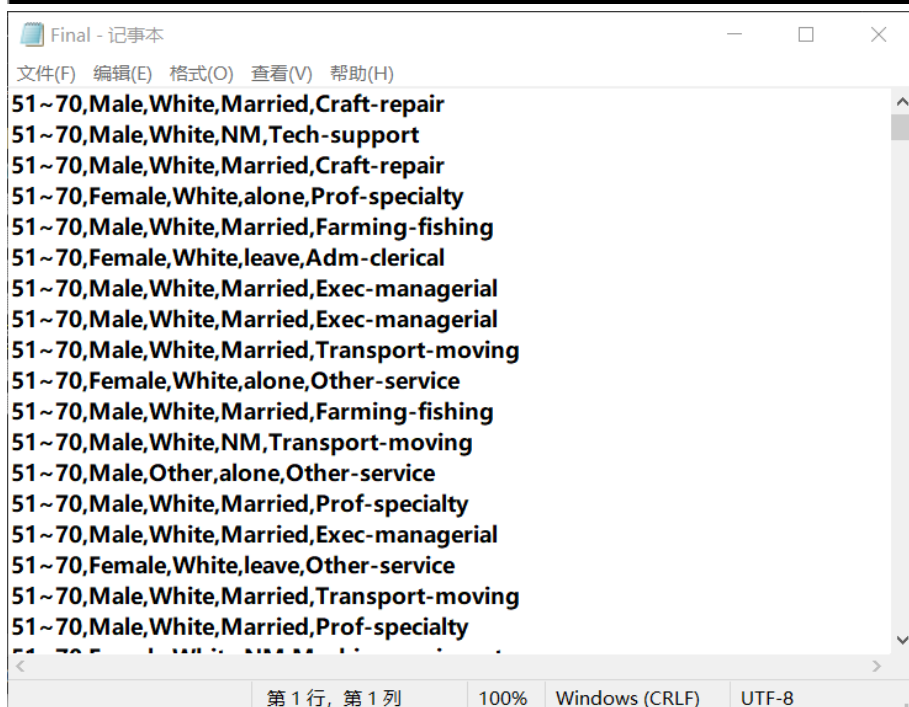
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

71~80,Male,Black,\*,Sales  
71~80,Male,White,\*,Craft-repair  
71~80,Male,White,\*,Exec-managerial  
71~80,Male,White,\*,Sales  
71~80,Male,White,\*,Prof-specialty  
71~80,Male,White,\*,Protective-serv  
71~80,Male,White,\*,Protective-serv  
71~80,Male,White,\*,Exec-managerial  
71~80,Male,Black,\*,Craft-repair  
71~80,Male,White,\*,Exec-managerial  
71~80,Male,White,\*,Exec-managerial  
71~80,Male,White,\*,Other-service  
71~80,Male,White,\*,Prof-specialty  
71~80,Male,White,\*,Transport-moving  
71~80,Male,White,\*,Exec-managerial  
71~80,Male,Black,\*,Protective-serv  
71~80,Male,Black,\*,Other-service  
71~80,Male,White,\*,Prof-specialty  
71~80,Male,White,\*,Prof-specialty

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

K=1000, maxsup=500 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Samarati
$ python main.py 1000 500
[3, 0, 0, 1]
LM=0.352966
Running time: 17.373523235321045 Seconds
```



Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

51~70,Male,White,Married,Craft-repair  
51~70,Male,White,NM,Tech-support  
51~70,Male,White,Married,Craft-repair  
51~70,Female,White,alone,Prof-specialty  
51~70,Male,White,Married,Farming-fishing  
51~70,Female,White,leave,Adm-clerical  
51~70,Male,White,Married,Exec-managerial  
51~70,Male,White,Married,Exec-managerial  
51~70,Male,White,Married,Transport-moving  
51~70,Female,White,alone,Other-service  
51~70,Male,White,Married,Farming-fishing  
51~70,Male,White,NM,Transport-moving  
51~70,Male,Other,alone,Other-service  
51~70,Male,White,Married,Prof-specialty  
51~70,Male,White,Married,Exec-managerial  
51~70,Female,White,leave,Other-service  
51~70,Male,White,Married,Transport-moving  
51~70,Male,White,Married,Prof-specialty

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

观察上面的测试结果可知：

当  $k$  不变,  $\maxsup$  变大时，运行时间变小，LM 变小（因为可以通过不发布数据来减小泛化深度，因为  $\maxsup$  变大到一定程度，对于不同泛化深度尝试次数减少，运行时间会减少）。

## 4, 讨论与总结

不同的  $k$  与  $\maxsup$  对算法输出和性能有影响。同时算法理论上来说可以得到很多解满足要求。而具体的运行时间在不同测试环境下均有一定偏差。

## 二 . Mondrian 算法

### 1 . 问题描述

使用 Mondrain 算法完成  $k$  匿名。

要求：

$QI=\{\text{age, education\_num}\}$ （数值型）， $S = \{\text{occupation}\}$ ;

输入：  $\text{data}, k, \maxSup$  ( $\text{data}$  是数据集，  $k$  是  $K$ -Anonymity 的参数，  $\maxSup$  表示最大 suppression 的个数);

输出： 匿名后的数据集。

评价指标：运行时间和 LM

### 2 . 程序使用指南以及程序分析

注：本程序也实现了附加的 Mondrian 算法实现非数值型的泛化

只需在修改如下代码  $QI\_columns$  处需要泛化的属性即可：

```

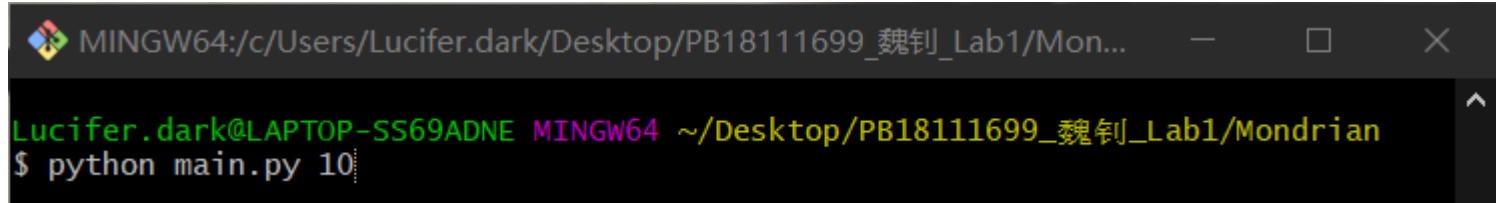
42 # QI 和 S
43 QI_columns = ['age', 'education_num', 'marital_status', 'sex', 'race']
44 S_columns = ['occupation', 'class']

```

这里对 age, education\_num, marital\_status, sex, race 进行了泛化, S 集为 occupation, class

在 Mondrian 文件夹下在终端执行 main.py 文件。

例如: K=10



```

MINGW64:/c/Users/Lucifer.dark/Desktop/PB18111699_魏钊_Lab1/Mon...
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Mondrian
$ python main.py 10

```

程序的输出为 Mondrian 文件下的 Final.txt(泛化后的数据), 和终端中打印的两行数据。

两行数据分别为:

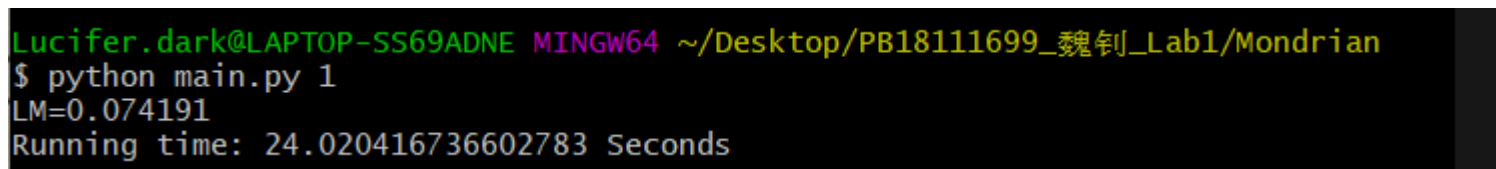
LM

运行时间

相关程序分析, 在代码注释中有详细的描述。

### 3. 实验结果分析

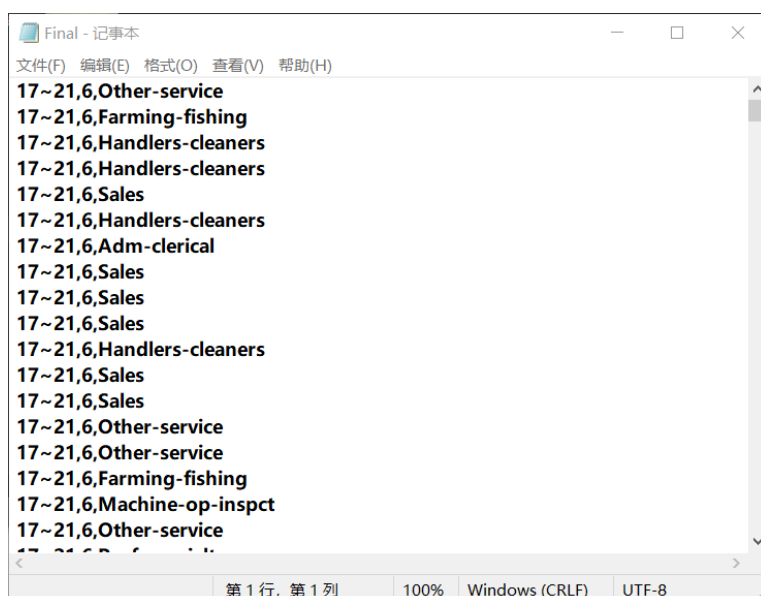
K=1 时:



```

Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Mondrian
$ python main.py 1
LM=0.074191
Running time: 24.020416736602783 Seconds

```



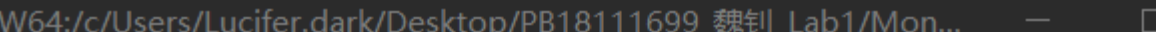
```

Final - 记事本
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)
17~21,6,Other-service
17~21,6,Farming-fishing
17~21,6,Handlers-cleaners
17~21,6,Handlers-cleaners
17~21,6,Sales
17~21,6,Handlers-cleaners
17~21,6,Adm-clerical
17~21,6,Sales
17~21,6,Sales
17~21,6,Sales
17~21,6,Handlers-cleaners
17~21,6,Sales
17~21,6,Sales
17~21,6,Other-service
17~21,6,Other-service
17~21,6,Farming-fishing
17~21,6,Machine-op-inspct
17~21,6,Other-service

```



K=10 时:



The screenshot shows a Windows command prompt window with a dark background. The title bar at the top reads "MINGW64:/c/Users/Lucifer.dark/Desktop/PB18111699\_魏钊\_Lab1/Mon...". The prompt shows the user "Lucifer.dark@LAPTOP-SS69ADNE" in a green color, followed by "MINGW64" in magenta and the current directory "~/Desktop/PB18111699\_魏钊\_Lab1/Mondrian" in yellow. The user has entered the command "\$ python main.py 10". The output of the command is displayed on the next two lines: "LM=0.077385" and "Running time: 23.725613117218018 Seconds".

```
MINGW64:/c/Users/Lucifer.dark/Desktop/PB18111699_魏钊_Lab1/Mon...  
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Mondrian  
$ python main.py 10  
LM=0.077385  
Running time: 23.725613117218018 Seconds
```

Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

17~21,6,Other-service  
 17~21,6,Farming-fishing  
 17~21,6,Handlers-cleaners  
 17~21,6,Handlers-cleaners  
 17~21,6,Sales  
 17~21,6,Handlers-cleaners  
 17~21,6,Adm-clerical  
 17~21,6,Sales  
 17~21,6,Sales  
 17~21,6,Sales  
 17~21,6,Handlers-cleaners  
 17~21,6,Sales  
 17~21,6,Sales  
 17~21,6,Other-service  
 17~21,6,Other-service  
 17~21,6,Farming-fishing  
 17~21,6,Machine-op-inspct  
 17~21,6,Other-service

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

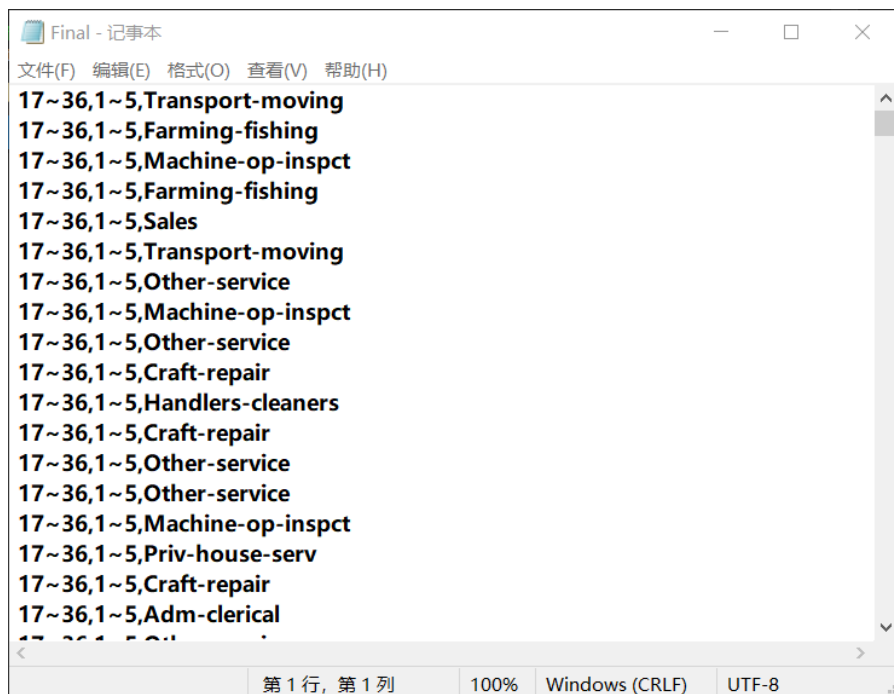
K=100 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Mondrian
$ python main.py 100
LM=0.111338
Running time: 23.77035164833069 Seconds
```



K=500 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Mondrian
$ python main.py 500
LM=0.233489
Running time: 25.07267737388611 Seconds
```



Final - 记事本

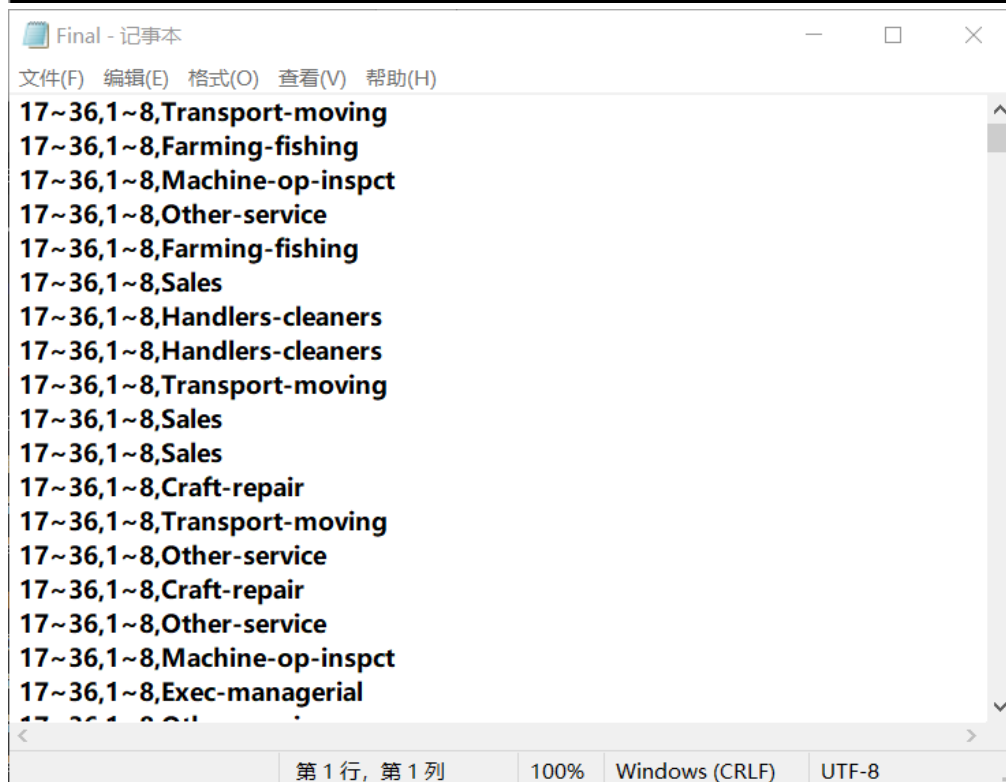
文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
17~36,1~5,Transport-moving
17~36,1~5,Farming-fishing
17~36,1~5,Machine-op-inspct
17~36,1~5,Farming-fishing
17~36,1~5,Sales
17~36,1~5,Transport-moving
17~36,1~5,Other-service
17~36,1~5,Machine-op-inspct
17~36,1~5,Other-service
17~36,1~5,Craft-repair
17~36,1~5,Handlers-cleaners
17~36,1~5,Craft-repair
17~36,1~5,Other-service
17~36,1~5,Other-service
17~36,1~5,Machine-op-inspct
17~36,1~5,Priv-house-serv
17~36,1~5,Craft-repair
17~36,1~5,Adm-clerical
```

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

K=1000 时:

```
Lucifer.dark@LAPTOP-SS69ADNE MINGW64 ~/Desktop/PB18111699_魏钊_Lab1/Mondrian
$ python main.py 1000
LM=0.399581
Running time: 25.791969537734985 Seconds
```



Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

```
17~36,1~8,Transport-moving
17~36,1~8,Farming-fishing
17~36,1~8,Machine-op-inspct
17~36,1~8,Other-service
17~36,1~8,Farming-fishing
17~36,1~8,Sales
17~36,1~8,Handlers-cleaners
17~36,1~8,Handlers-cleaners
17~36,1~8,Transport-moving
17~36,1~8,Sales
17~36,1~8,Sales
17~36,1~8,Craft-repair
17~36,1~8,Transport-moving
17~36,1~8,Other-service
17~36,1~8,Craft-repair
17~36,1~8,Other-service
17~36,1~8,Machine-op-inspct
17~36,1~8,Exec-managerial
```

第 1 行, 第 1 列 100% Windows (CRLF) UTF-8

可以看到当 k 增加时，LM 变大。因为 k 越大，二分的层数就越小，而 LM (loss 就越大。

下面进行选做的测试：

QI 和 S 集如下：

```
# QI 和 S
QI_columns = ['age', 'education_num', 'marital_status', 'sex', 'race']
S_columns = ['occupation']
```

输出格式如下图所示。

K=1000 时：

Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Other-service
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Machine-op-inspct
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Adm-clerical
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Other-service
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Machine-op-inspct
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Other-service
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Other-service
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Transport-moving
37~90,1~9,Separated or Never-married or Widowed or Married-AF-spouse ,Female or Male ,Black or White or Other or Asian-Pac-Islander or Amer-Indian-Eskimo ,Other-service

K=10 时：

Final - 记事本

文件(F) 编辑(E) 格式(O) 查看(V) 帮助(H)

22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Prof-specialty
22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Adm-clerical
22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Sales
22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Other-service
22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Sales
22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Tech-support
22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Adm-clerical
22~36,10~14,Married-spouse-absent or Separated or Widowed ,Female ,Asian-Pac-Islander or Other or Amer-Indian-Eskimo ,Prof-specialty

### 三 . 选做部分

1 . Samarati 算法可能会有很多解满足要求, 调研并探究如何选择输出保证结果的可用性尽可能大, 说说你的启发。

可以注意到在 Samarati 算法中当取得中间高度的向量组时, 不同的枚举方式可以得到不同的向量组顺序, 因而得到的第一个满足  $k$  匿名的向量组有所不同, 所以导致多解性。

通过对比不同枚举顺序导致的发布的数据的差别可以看出, 当我们枚举向量组时, 可以优先枚举泛化层数小的属性, 这样可用性, 隐私和 LM 可以得到一个更好的平衡, 数据可用性也更大。

抛开数据属性本身的泛化层数和属性范围来说, 可以根据实际情况以及不同属性的可用性角度来控制枚举的顺序。比如: 对于 QI 中可能发布后比较敏感的属性, 可以最后枚举 (这样泛化程度一般更高) 或者对于 QI 中可能一些需要精确发布的属性, 可以优先枚举。从而可以通过控制枚举顺序来得到不同的发布数据, 在满足同样的  $k$  匿名的前提下, 数据可用性变大。

2.Mondrian 算法处理 categorical, 已经在 Mondrian 部分说明过了。