

FIES Data Anonymization Usecase1

Abdellahi El Moustapha

2025-03-30

Introduction

This document demonstrates how to transform the 2006 Family Income and Expenditure Survey (FIES) data into a safe microdata file using the `sdcMicro` package. We will: - Pre-process and inspect the data

- Create an `sdcMicro` object
- Evaluate disclosure risk
- Recode key variables
- Apply local suppression, microaggregation, and PRAM
- Export the anonymized data

1. Data Import and Pre-processing

```
# Load required libraries
library(sdcMicro)
library(readxl)
library(sdcTable)

## Loading required package: Rcpp
## Loading required package: sdcHierarchies
## Loading required package: shinythemes
## Package "sdcHierarchies" 0.21.0 has been loaded.
## Package sdcTable 0.32.7 has been loaded!

# Read the FIES 2006 data from a CSV file (adjust the file path as needed)
fies06 <- read.csv("data/FIES06.csv")

# Convert all column names to uppercase
names(fies06) <- toupper(names(fies06))

# Verify the structure
str(fies06)

## 'data.frame':    38483 obs. of  721 variables:
##  $ W_REGN      : int  14 14 14 14 14 14 14 14 14 14 ...
##  $ W_URB2      : int   2 2 2 2 2 2 2 2 2 2 ...
##  $ W_STR2      : int 1177 1177 1177 1177 1177 1177 1177 1177 1177 1177 ...
##  $ W_PSU       : int   1 1 1 1 1 1 1 1 1 1 ...
##  $ PUFHHSEQNO  : int   1 2 3 4 5 6 7 8 9 10 ...
##  $ W_REP       : int   3 3 3 3 3 3 3 3 3 3 ...
```

```

## $ W_ROUTE      : chr  "07" "07" "07" "07" ...
## $ W_INT_STS     : int   1 1 1 1 1 1 1 1 1 ...
## $ W_LINE_NO     : int   2 1 2 1 5 1 1 1 2 2 ...
## $ W_NO_HH       : int   1 1 1 1 1 1 1 1 1 ...
## $ FSIZE         : num   6 5 8 7.5 5 4 4 5 8.5 4 ...
## $ RFACT         : num   185 185 185 185 185 ...
## $ Z2011_H_SEX   : int   1 1 1 1 1 1 1 1 1 ...
## $ Z2021_H_AGE    : int   30 51 51 24 59 78 54 38 33 26 ...
## $ Z2031_H_MS     : int   2 2 2 2 4 2 2 2 2 ...
## $ Z2041_H_EDUC   : int   4 4 4 2 1 1 4 5 3 1 ...
## $ Z2051_H_HAS_JOB : int   1 1 2 1 1 1 1 1 1 ...
## $ Z2061_H_OCCUP  : int   7142 6111 NA 6111 7123 6111 6111 7123 7123 6111 ...
## $ Z2071_H_KB     : int   4541 111 NA 111 4520 111 111 4520 4520 111 ...
## $ Z2081_H_CW     : int   1 4 NA 3 1 3 3 1 1 3 ...
## $ Z2091_HHLD_TYPE : int   1 1 2 1 2 2 1 1 2 1 ...
## $ Z2101_TOT_MEM  : int   6 5 8 8 6 3 4 5 9 4 ...
## $ Z2111_M_LESS_1 : int   NA NA NA 1 1 NA NA NA NA NA ...
## $ Z2121_M_LESS_7 : int   2 NA NA 2 NA NA NA 2 1 2 ...
## $ Z2131_M_LESS_15 : int  2 2 1 3 NA NA NA 2 3 NA ...
## $ Z2141_M_LESS_25 : int   NA 2 3 1 2 NA 2 NA 2 NA ...
## $ Z2151_M_25_OVER : int   2 1 2 1 3 1 2 1 3 2 ...
## $ Z2152_M_60_OVER : int   NA NA 2 NA NA 2 NA NA NA NA ...
## $ Z2161_M_TOT_NREL : int   NA NA NA NA NA NA NA NA NA NA ...
## $ Z2171_M_TOT_EMP : int   2 1 2 1 2 1 1 1 4 1 ...
## $ Z2181_WIFE_EMP  : int   1 1 1 2 3 2 2 1 2 2 ...
## $ B4011_BLDG_TYPE : int   1 1 1 1 1 1 1 1 1 1 ...
## $ B4021_ROOF     : int   4 1 1 1 1 1 1 1 1 1 ...
## $ B4031_WALLS     : int   2 1 1 5 1 1 1 1 1 5 ...
## $ B4041_TENURE    : int   4 1 1 1 1 1 1 1 1 4 ...
## $ B4042_TENURE_IND : logi  NA NA NA NA NA NA ...
## $ B4043_HOUSE_RENT : int   1200 0 0 0 0 0 0 0 0 2400 ...
## $ B4053_LOT_RENT  : int   2400 12000 30000 3600 24000 30000 12000 18000 12000 2400 ...
## $ B4081_HSE_ALTERTN : int  2 2 2 2 2 2 2 2 2 ...
## $ B5012_OTH_HOUSE : int   0 0 0 0 0 0 0 0 0 ...
## $ B5021_TOILET    : int   1 1 1 5 1 1 1 1 1 ...
## $ B5031_ELECTRIC   : int  2 1 1 2 1 1 1 1 2 ...
## $ B5041_WATER      : int   4 1 1 4 1 1 1 3 1 ...
## $ B5042_DISTANCE  : int   2 0 0 100 5 0 0 0 1 75 ...
## $ B5051_W_RADIO    : int   2 1 1 1 1 2 1 1 2 ...
## $ B5052_N_RADIO    : int   NA 1 1 1 1 NA 1 1 NA NA ...
## $ B5061_W_TV       : int   2 2 1 2 1 1 1 1 2 ...
## $ B5062_N_TV       : int   NA NA 1 NA 1 1 1 1 NA NA ...
## $ B5071_VTR        : int   2 2 1 2 2 2 1 1 2 ...
## $ B5072_N_VTR      : int   NA NA 1 NA NA NA 1 1 NA NA ...
## $ B5081_W_STEREO   : int   2 1 2 2 2 2 2 2 2 ...
## $ B5082_N_STEREO   : int   NA 1 NA NA NA NA NA NA NA ...
## $ B5091_W_REF      : int   2 1 1 2 2 2 2 2 2 ...
## $ B5092_N_REF      : int   NA 1 1 NA NA NA NA NA NA ...
## $ B5101_W_WASH     : int   2 2 1 2 2 2 2 2 2 ...
## $ B5102_N_WASH     : int   NA NA 1 NA NA NA NA NA NA ...
## $ B5111_W_AIRCON   : int   2 2 2 2 2 2 2 2 2 ...
## $ B5112_N_AIRCON   : int   NA NA NA NA NA NA NA NA NA ...
## $ B5121_W_SALASET  : int   2 2 2 2 2 2 2 2 2 ...
## $ B5122_N_SALASET  : int   NA NA NA NA NA NA NA NA NA ...

```

```
## $ B5131_W_DINING : int 2 2 2 2 2 2 2 2 2 ...
## $ B5132_N_DINING : int NA NA NA NA NA NA NA NA NA ...
## $ B5141_W_CAR : int 2 2 2 2 2 2 2 2 2 ...
## $ B5142_N_CAR : int NA NA NA NA NA NA NA NA NA ...
## $ B5151_W_PHONE : int 2 1 1 2 2 2 2 1 2 2 ...
## $ B5152_N_PHONE : int NA 1 1 NA NA NA NA 1 NA NA ...
## $ B5161_W_PC : int 2 2 2 2 2 2 2 2 2 ...
## $ B5162_N_PC : int NA NA NA NA NA NA NA NA NA ...
## $ B5171_W_OVEN : int 2 2 2 2 2 2 2 2 2 ...
## $ B5172_N_OVEN : int NA NA NA NA NA NA NA NA NA ...
## $ B5181_W_MOTOR : int 2 1 1 2 2 2 2 2 2 ...
## $ B5182_N_MOTOR : int NA 1 1 NA NA NA NA NA NA ...
## $ TOTEX : int 53142 189105 200101 47630 95456 97887 116809 108273 110081 43097 ...
## $ CTEXP : int 44759 172913 152498 31612 61864 56954 88083 85737 79555 28175 ...
## $ KTEXP : int 8383 16192 47603 16018 33592 40933 28726 22536 30526 14922 ...
## $ FOOD : int 35267 107848 107151 29769 42422 43947 59599 67751 75314 26279 ...
## $ CFOOD : int 33796 107602 95206 21359 37557 35736 46026 66911 60668 19326 ...
## $ KFOOD : int 1471 246 11945 8410 4865 8211 13573 840 14646 6953 ...
## $ CREAL : int 13863 26979 32634 14612 16892 16770 19919 16804 29459 10514 ...
## $ CRAL : int 13863 26979 24409 11912 15492 13060 13219 16804 26039 7976 ...
## $ KRAL : int 0 0 8225 2700 1400 3710 6700 0 3420 2538 ...
## $ TRICE : int 9919 18110 24804 13780 13780 13286 17043 12792 25935 9646 ...
## $ CRICE : int 9919 18110 16579 11080 12380 9576 10343 12792 22515 7108 ...
## $ KRICE : int 0 0 8225 2700 1400 3710 6700 0 3420 2538 ...
## $ TSPEC : int 0 0 0 0 0 0 0 0 0 0 ...
## $ CSPEC : int 0 0 0 0 0 0 0 0 0 0 ...
## $ KSPEC : int 0 0 0 0 0 0 0 0 0 0 ...
## $ TORDC : int 9919 18110 24804 13780 13780 13286 17043 12792 25935 9646 ...
## $ CORDC : int 9919 18110 16579 11080 12380 9576 10343 12792 22515 7108 ...
## $ KORDC : int 0 0 8225 2700 1400 3710 6700 0 3420 2538 ...
## $ TNFAR : int 0 0 0 0 0 0 0 0 0 0 ...
## $ CNFAR : int 0 0 0 0 0 0 0 0 0 0 ...
## $ KNFAR : int 0 0 0 0 0 0 0 0 0 0 ...
## $ TOTHR : int 0 0 0 0 0 0 0 0 0 0 ...
## $ COTHR : int 0 0 0 0 0 0 0 0 0 0 ...
## $ KOTHR : int 0 0 0 0 0 0 0 0 0 0 ...
## $ TCORN : int 100 120 0 0 0 0 0 0 0 0 ...
## $ CCORN : int 100 120 0 0 0 0 0 0 0 0 ...
## $ KCORN : int 0 0 0 0 0 0 0 0 0 0 ...
## [list output truncated]
```

COMMENT:

The dataset contains 38,483 observations and 721 variables. Column names are now fully capitalized, as seen in W_REGN, W_URB2, etc. This step ensures consistency for referencing variables in `sdcmicro`, which is case-sensitive and requires exact variable names.

2. Create `sdcmicro` Object

```
# Create an sdcmicro object with key variables and numerical variables,
# and specify the sampling weight variable ('RFACT').
# Note: All variable names are now in uppercase.
sdc <- createSdcObj(fies06,
```

```

keyVars = c('W_REGN', 'Z2011_H_SEX', 'Z2021_H_AGE', 'Z2041_H_EDUC'),
numVars = c('WAGES', 'WSAG', 'WSNAG', 'PNSNS', 'TOINC'),
weightVar = 'RFACT')

# Display a summary of the sdcMicro object (includes frequency counts and anonymity info)
show(sdc)

## The input dataset consists of 38483 rows and 721 variables.
## --> Categorical key variables: W_REGN, Z2011_H_SEX, Z2021_H_AGE, Z2041_H_EDUC
## --> Numerical key variables: WAGES, WSAG, WSNAG, PNSNS, TOINC
## --> Weight variable: RFACT
## -----

## Information on categorical key variables:
##
## Reported is the number, mean size and size of the smallest category >0 for recoded variables.
## In parenthesis, the same statistics are shown for the unmodified data.
## Note: NA (missings) are counted as separate categories!

## Key Variable Number of categories      Mean size
##      <char>          <char> <char>      <char>      <char>
##      W_REGN              17  (17)  2263.706  (2263.706)
##      Z2011_H_SEX          2   (2) 19241.500 (19241.500)
##      Z2021_H_AGE         86  (86)   447.477  (447.477)
##      Z2041_H_EDUC         25  (25)  1539.320  (1539.320)
## Size of smallest (>0)
##      <char> <char>
##      1536 (1536)
##      6939 (6939)
##      1    (1)
##      1    (1)
## -----

## Infos on 2/3-Anonymity:
##
## Number of observations violating
## - 2-anonymity: 4506 (11.709%)
## - 3-anonymity: 8068 (20.965%)
## - 5-anonymity: 13753 (35.738%)
## -----

## Numerical key variables: WAGES, WSAG, WSNAG, PNSNS, TOINC
##
## Disclosure risk is currently between [0.00%; 100.00%]
##
## Current Information Loss:
## - IL1: 0.00
## - Difference of Eigenvalues: 0.000%
## -----

```

COMMENT:

This output shows violations of k-anonymity: - 4,506 records violate 2-anonymity (more easily re-identifiable) - 8,068 violate 3-anonymity, and 13,753 violate 5-anonymity This highlights the high re-identification risk in

the original dataset.

Numerical key variables include: WAGES, WSAG, WSNAG, PNSNS, TOINC

Disclosure risk bounds are very wide (0%–100%), indicating the need for protection.

Information loss (IL1 and eigenvalue difference) is currently 0.00, which is expected since no anonymization techniques have been applied yet.

3. Frequency Counts and Disclosure Risk

```
# Print risk measures: expected number of re-identifications, etc.
print(sdc, "risk")

## Risk measures:
##
## Number of observations with higher risk than the main part of the data: 0
## Expected number of re-identifications: 91.31 (0.24 %)
```

COMMENT:

This output summarizes the disclosure risk in the dataset. No records have higher risk than the main part of the population, which is reassuring.

The expected number of re-identifications is approximately 91 out of 38,483 records (0.24%). This level of risk is relatively low, but it still justifies the need for anonymization techniques to further minimize potential disclosure.

4. Global Risk Estimation (Log-linear Model)

```
# If available, include additional variables (e.g., 'Z2031_H_MS') in the risk model.
# Here we check if 'Z2031_H_MS' exists; if not, we skip this step.
if("Z2031_H_MS" %in% names(fies06)) {
  form <- ~W_REGN + Z2011_H_SEX + Z2021_H_AGE + Z2041_H_EDUC + Z2031_H_MS
  sdc <- modRisk(sdc, formulaM = form)
  # Display the estimated model and global risk measures
  print(slot(sdc, "risk")$model)
} else {
  cat("Variable 'Z2031_H_MS' not found; skipping modRisk model.\n")
}
```

```
## The estimated model (using method 'default') was:
## ~ W_REGN + Z2011_H_SEX + Z2021_H_AGE + Z2041_H_EDUC + Z2031_H_MS
## global risk-measures:
## Risk-Measure 1: 0.433 (43.300 %)
## Risk-Measure 2: 0.485 (48.529 %)
```

COMMENT:

The risk measures indicate the level of disclosure risk in the dataset. In this case, the risk measures are relatively high, suggesting that there is a significant risk of re-identification. This highlights the need for further anonymization techniques to minimize potential disclosure risks.

5. Recoding Key Variables

```
# Recode the age variable into 10-year intervals to reduce key uniqueness
sdc <- globalRecode(sdc, column = "Z2021_H_AGE",
                    breaks = seq(9, 99, 10), labels = 1:9)

# Recode the education variable by combining categories
sdc <- globalRecode(sdc, column = "Z2041_H_EDUC",
                    breaks = c(-100, 59, 69, 180), labels = c(0, 6, 7))

# Display k-anonymity status after recoding
print(sdc, "kAnon")
```

```
## Infos on 2/3-Anonymity:
##
## Number of observations violating
## - 2-anonymity: 84 (0.218%) | in original data: 4506 (11.709%)
## - 3-anonymity: 158 (0.411%) | in original data: 8068 (20.965%)
## - 5-anonymity: 334 (0.868%) | in original data: 13753 (35.738%)
##
## -----
```

COMMENT:

The recoding of key variables has significantly reduced the number of observations violating k-anonymity, which indicates improved anonymity in the dataset. However, there are still some observations violating 2-anonymity, 3-anonymity, and 5-anonymity, suggesting that further anonymization techniques may be necessary to fully protect individual privacy.

6. Local Suppression

```
# Apply local suppression to achieve 3-anonymity on the key variables
sdc <- localSuppression(sdc, k = 3)

# Print details on local suppression and current k-anonymity status
print(sdc, "ls")
```

```
## Local suppression:
##      KeyVar      | Suppressions (#)      | Suppressions (%)
##      <char> <char>      <int> <char>      <char>
##      W_REGN      |      155      |      0.403
##      Z2011_H_SEX  |      0      |      0.000
##      Z2021_H_AGE  |      4      |      0.010
##      Z2041_H_EDUC |      0      |      0.000
## -----
```

```
print(sdc, "kAnon")
```

```
## Infos on 2/3-Anonymity:
##
## Number of observations violating
## - 2-anonymity: 0 (0.000%) | in original data: 4506 (11.709%)
## - 3-anonymity: 0 (0.000%) | in original data: 8068 (20.965%)
```

```
## - 5-anonymity: 29 (0.075%) | in original data: 13753 (35.738%)
##
## -----
```

COMMENT:

-The local suppression process has resulted in a small number of suppressions for the key variables. This indicates that the technique has been applied to reduce the uniqueness of the key variables, thereby enhancing anonymity. The k-anonymity status shows the number of observations violating different levels of anonymity after the suppression process. -After applying local suppression, the dataset has achieved 3-anonymity with no observations violating 2-anonymity or 3-anonymity. However, there are still 29 observations violating 5-anonymity. This indicates that while the local suppression technique has improved the anonymity of the dataset, further anonymization techniques may be necessary to fully protect individual privacy, especially for higher levels of k-anonymity.

7. Microaggregation for Numerical Variables

```
# Apply microaggregation (using the MDAV method) to reduce risk on numerical variables.
# This aggregates values into groups of similar observations.
sdc <- microaggregation(sdc)
```

```
# Display numerical risk measures after microaggregation
print(sdc, "numrisk")
```

```
## Numerical key variables: WAGES, WSAG, WSNAG, PNSNS, TOINC
##
## Disclosure risk (~100.00% in original data):
##   modified data: [0.00%; 91.51%]
##
## Current Information Loss in modified data (0.00% in original data):
##   IL1: 315480.19
##   Difference of Eigenvalues: -72.770%
## -----
```

COMMENT:

Disclosure Risk: The disclosure risk in the original data was 100%, indicating a high risk of re-identification. After applying microaggregation, the disclosure risk has been significantly reduced, ranging from 0.00% to 91.51%. This indicates that the technique has effectively reduced the risk of re-identification. Information Loss: The information loss (IL1) in the modified data is 315480.19, which is a measure of how much information has been lost due to the aggregation process. The difference of eigenvalues is -72.770%, indicating a reduction in the variability of the data.

Overall, the application of microaggregation has successfully reduced the disclosure risk while introducing some information loss. This technique helps in balancing the need for data privacy with the utility of the dataset.

8. Post-Randomization (PRAM) for Categorical Variables

```
# Ensure 'W_REGN' is a factor in both the original data and the manipulated key variables
sdc@origData$W_REGN <- as.factor(sdc@origData$W_REGN)
sdc@manipKeyVars$W_REGN <- as.factor(sdc@manipKeyVars$W_REGN)
```

```
# Now apply PRAM to the 'W_REGN' variable using the correct argument 'variables'
sdc <- pram(sdc, variables = "W_REGN")
```

```
## Warning in pramX(obj = obj, variables = variables, strata_variables = strata_variables, : If pram is
```

```
# Show the first few lines of the PRAM transition summary
```

```
head(get.sdcMicroObj(sdc, "pram"))
```

```
## $params
```

```
## $params$W_REGN
```

```
## $params$W_REGN$Rs
```

```
##           1           2           3           5           6           7
## 1  0.887664142 0.003592647 0.010475778 0.009201670 0.005576666 0.008532183
## 2  0.004287937 0.931688775 0.007472629 0.006292908 0.002414402 0.005535437
## 3  0.007587187 0.004534549 0.885096528 0.009136055 0.006221948 0.008618013
## 5  0.009259258 0.005305510 0.012693276 0.858751016 0.007472287 0.010590184
## 6  0.004632131 0.001680281 0.007135725 0.006168084 0.927827489 0.005608958
## 7  0.007691608 0.004180951 0.010726785 0.009487482 0.006087416 0.882380894
## 8  0.007943785 0.003775010 0.011516529 0.010093132 0.005999616 0.009333667
## 9  0.010783540 0.005670187 0.015193269 0.013405199 0.008433159 0.012496022
## 10 0.007532221 0.003066299 0.011333808 0.009848592 0.005417215 0.009013956
## 11 0.007191359 0.003232988 0.010574500 0.009237193 0.005333719 0.008508497
## 12 0.005073230 0.001213617 0.008320769 0.007095422 0.003197976 0.006343196
## 13 0.008006789 0.005471552 0.010263413 0.009269030 0.006929217 0.008875487
## 14 0.005603332 0.000977425 0.009483043 0.008033756 0.003340020 0.007121917
## 15 0.006956923 0.002425253 0.010796357 0.009317094 0.004788141 0.008455407
## 16 0.007086291 0.002519016 0.010957864 0.009463950 0.004902934 0.008597100
## 41 0.004880814 0.002432924 0.006984427 0.006139847 0.003746340 0.005698544
## 42 0.008898979 0.004159296 0.012957502 0.011344552 0.006684181 0.010478224
##           8           9          10          11          12          13
## 1  0.006828620 0.007468475 0.005738676 0.006469347 0.004313710 0.01581083
## 2  0.003873088 0.004687074 0.002788283 0.003471262 0.001231635 0.01289557
## 3  0.007170040 0.007621076 0.006254016 0.006889756 0.005124180 0.01467853
## 5  0.008730537 0.009342291 0.007550441 0.008361787 0.006070915 0.01841790
## 6  0.004283859 0.004851400 0.003428238 0.003985528 0.002258644 0.01136545
## 7  0.007232938 0.007801882 0.006191003 0.006900163 0.004862185 0.01579957
## 8  0.877293533 0.008156770 0.006199378 0.007016202 0.004587557 0.01751720
## 9  0.010124022 0.833815982 0.008611098 0.009635094 0.006663693 0.02257376
## 10 0.006994633 0.007827814 0.882784369 0.006548572 0.003939729 0.01774323
## 11 0.006704371 0.007417832 0.005546075 0.888681926 0.004005591 0.01626549
## 12 0.004637891 0.005427742 0.003530112 0.004237890 0.919981693 0.01383216
## 13 0.007625596 0.007917323 0.006845813 0.007410036 0.005956074 0.88126006
## 14 0.005091281 0.006061131 0.003768883 0.004607046 0.001888810 0.01607941
## 15 0.006425401 0.007304186 0.005113575 0.005965846 0.003313622 0.01728683
## 16 0.006549071 0.007431130 0.005226104 0.006086660 0.003414138 0.01750078
## 41 0.004570822 0.004990966 0.003849687 0.004333918 0.002909394 0.01051219
## 42 0.008309998 0.009150283 0.006919861 0.007845730 0.005083523 0.01977738
##           14          15          16          41          42
## 1  0.0037886586 0.005077843 0.005137640 0.007808001 0.006515112
## 2  0.0007887809 0.002112773 0.002179763 0.004645260 0.003634422
## 3  0.0046438838 0.005707341 0.005753937 0.008092318 0.006870644
## 5  0.0054659710 0.006843087 0.006904410 0.009883590 0.008357542
## 6  0.0018758344 0.002902921 0.002952616 0.004978070 0.004064772
## 7  0.0043410321 0.005563570 0.005618929 0.008218045 0.006915544
```



```

## 8 0.0040046151 0.005455779 0.005523557 0.008506218 0.007077455
## 9 0.0059172815 0.007697740 0.007779092 0.011528204 0.009672653
## 10 0.0033447463 0.004898894 0.004973179 0.008083221 0.006649522
## 11 0.0034626788 0.004840437 0.004905397 0.007706883 0.006385068
## 12 0.0015019692 0.002844452 0.002911115 0.005473734 0.004377036
## 13 0.0055057065 0.006389709 0.006425482 0.008516178 0.007332535
## 14 0.9111186088 0.002947674 0.003028391 0.006058210 0.004791060
## 15 0.0027305905 0.891163972 0.004383855 0.007479876 0.006093077
## 16 0.0028242713 0.004413403 0.889197810 0.007617290 0.006212190
## 41 0.0025605651 0.003412784 0.003452212 0.925162488 0.004362072
## 42 0.0044247529 0.006074591 0.006151878 0.009531444 0.862207827
##
## $params$W_REGN$pd
## [1] 0.8
##
## $params$W_REGN$alpha
## [1] 0.5
##
##
## $transitions
## $transitions$W_REGN
##      transition Frequency
##      <char>      <int>
## 1: 1 --> 1      1985
## 2: 1 --> 10      15
## 3: 1 --> 11      16
## 4: 1 --> 12      33
## 5: 1 --> 13      10
## ---
## 285: 9 --> 6      19
## 286: 9 --> 7       8
## 287: 9 --> 8      12
## 288: 9 --> 9     1528
## 289: NA --> NA     155
##
##
## $comparison
## $comparison$W_REGN
##
##      W_REGN      1      2      3      5      6      7
##      <char> <char> <char> <char> <char> <char> <char>
## 1:      Original Frequencies    2251    1886    3108    2237    2710    2497
## 2: Frequencies after Perturbation    2232    1904    3091    2296    2687    2490
##      8      9      10      11      12      13      14      15      16      41      42
##      <char> <char> <char> <char> <char> <char> <char> <char> <char> <char> <char>
## 1:    1935    1559    1715    2025    1914    4445    1522    1643    1632    3601    1648
## 2:    1924    1576    1721    2016    1912    4433    1537    1597    1640    3627    1645
##      NA
##      <char>
## 1:    155
## 2:    155
##
##
## $summary

```

```
## variable nrChanges percChanges
## 1 W_REGN 4107 10.67
```

COMMENT:

The application of PRAM to the W_REGN variable has introduced random noise, transitioning the original values to different values. This process helps in reducing the uniqueness of the data, thereby enhancing anonymity. The transition summary provides insights into how frequently each original value has been transitioned to other values, ensuring that the data remains useful while protecting individual privacy

9. Final Risk and Utility Evaluation

```
# Final Risk Evaluation
# Print the final risk measures after all anonymization steps
print(sdc, "risk")
```

```
## Risk measures:
##
## Number of observations with higher risk than the main part of the data:
## in modified data: 0
## in original data: 0
## Expected number of re-identifications:
## in modified data: 1.09 (0.00 %)
## in original data: 91.31 (0.24 %)
```

COMMENT:

Number of Observations with Higher Risk:

The number of observations with higher risk than the main part of the data is 0 in both the modified and original datasets. This indicates that the anonymization techniques have been effective in reducing the risk of re-identification. Expected Number of Re-identifications:

The expected number of re-identifications has been significantly reduced from 91.31 (0.24%) in the original data to 1.10 (0.00%) in the modified data. This demonstrates the effectiveness of the anonymization techniques in protecting individual privacy. Overall, the final risk evaluation shows that the anonymization techniques have successfully reduced the risk of re-identification while maintaining the utility of the dataset. This balance is crucial for ensuring that the data remains useful for analysis while protecting individual privacy.

```
# Evaluate data utility and information loss after the modifications
# Print the utility slot directly
print(sdc@utility)
```

```
## $il1
## [1] 315480.2
##
## $il1s
## [1] 1433.673
##
## $eigen
## [1] -0.7276922
```

COMMENT:

Information Loss (il1):

The information loss (il1) is 315480.2, which indicates the extent to which the data has been altered due to the anonymization techniques. This value suggests that there has been a noticeable amount of information loss. Standardized Information Loss (il1s):

The standardized information loss (il1s) is 1433.673. This standardized measure helps in understanding the relative impact of the modifications on the data's utility. Difference of Eigenvalues (eigen):

The difference of eigenvalues is -0.7276922, indicating a reduction in the variability of the data. This suggests that the modifications have made the data more uniform, which can be beneficial for anonymity but may also reduce the data's utility for certain types of analysis. Overall, the metrics indicate that while the anonymization techniques have introduced some information loss, they have also reduced the variability of the data. This balance is crucial for ensuring that the data remains useful for analysis while protecting individual privacy.

10. Export Anonymized Data

```
# Export Anonymized Data
# Extract the final anonymized microdata
final_data <- extractManipData(sdc)
head(final_data)
```

```
##   W_REGN W_URB2 W_STR2 W_PSU PUFHHSEQNO W_REP W_ROUTE W_INT_STS W_LINE_NO
## 1     14     2   1177     1         1     3      07         1         2
## 2     14     2   1177     1         2     3      07         1         1
## 3     14     2   1177     1         3     3      07         1         2
## 4     14     2   1177     1         4     3      07         1         1
## 5     11     2   1177     1         5     3      07         1         5
## 6     14     2   1177     1         6     3      07         1         1
##   W_NO_HH FSIZE   RFACT Z2011_H_SEX Z2021_H_AGE Z2031_H_MS Z2041_H_EDUC
## 1       1   6.0 185.4394           1           3           2           1
## 2       1   5.0 185.4394           1           5           2           1
## 3       1   8.0 185.4394           1           5           2           1
## 4       1   7.5 185.4394           1           2           2           1
## 5       1   5.0 185.4394           1           5           4           1
## 6       1   4.0 185.4394           1           7           2           1
##   Z2051_H_HAS_JOB Z2061_H_OCCUP Z2071_H_KB Z2081_H_CW Z2091_H_HLD_TYPE
## 1                1           7142       4541           1           1
## 2                1           6111          111           4           1
## 3                2              NA          NA          NA           2
## 4                1           6111          111           3           1
## 5                1           7123       4520           1           2
## 6                1           6111          111           3           2
##   Z2101_TOT_MEM Z2111_M_LESS_1 Z2121_M_LESS_7 Z2131_M_LESS_15 Z2141_M_LESS_25
## 1              6              NA              2              2              NA
## 2              5              NA              NA              2              2
## 3              8              NA              NA              1              3
## 4              8              1              2              3              1
## 5              6              1              NA              NA              2
## 6              3              NA              NA              NA              NA
##   Z2151_M_25_OVER Z2152_M_60_OVER Z2161_M_TOT_NREL Z2171_M_TOT_EMP
## 1                2              NA              NA              2
## 2                1              NA              NA              1
## 3                2              2              NA              2
## 4                1              NA              NA              1
```

## 5	3	NA	NA	2		
## 6	1	2	NA	1		
##	Z2181_WIFE_EMP	B4011_BLDG_TYPE	B4021_ROOF	B4031_WALLS	B4041_TENURE	
## 1	1	1	4	2	4	
## 2	1	1	1	1	1	
## 3	1	1	1	1	1	
## 4	2	1	1	5	1	
## 5	3	1	1	1	1	
## 6	2	1	1	1	1	
##	B4042_TENURE_IND	B4043_HOUSE_RENT	B4053_LOT_RENT	B4081_HSE_ALTERN		
## 1	NA	1200	2400		2	
## 2	NA	0	12000		2	
## 3	NA	0	30000		2	
## 4	NA	0	3600		2	
## 5	NA	0	24000		2	
## 6	NA	0	30000		2	
##	B5012_OTH_HOUSE	B5021_TOILET	B5031_ELECTRIC	B5041_WATER	B5042_DISTANCE	
## 1	0	1	2	4	2	
## 2	0	1	1	1	0	
## 3	0	1	1	1	0	
## 4	0	5	2	4	100	
## 5	0	1	1	1	5	
## 6	0	1	1	1	0	
##	B5051_W_RADIO	B5052_N_RADIO	B5061_W_TV	B5062_N_TV	B5071_VTR	B5072_N_VTR
## 1	2	NA	2	NA	2	NA
## 2	1	1	2	NA	2	NA
## 3	1	1	1	1	1	1
## 4	1	1	2	NA	2	NA
## 5	1	1	1	1	2	NA
## 6	2	NA	1	1	2	NA
##	B5081_W_STEREO	B5082_N_STEREO	B5091_W_REF	B5092_N_REF	B5101_W_WASH	
## 1	2	NA	2	NA	2	
## 2	1	1	1	1	2	
## 3	2	NA	1	1	1	
## 4	2	NA	2	NA	2	
## 5	2	NA	2	NA	2	
## 6	2	NA	2	NA	2	
##	B5102_N_WASH	B5111_W_AIRCON	B5112_N_AIRCON	B5121_W_SALASET	B5122_N_SALASET	
## 1	NA	2	NA	2	NA	
## 2	NA	2	NA	2	NA	
## 3	1	2	NA	2	NA	
## 4	NA	2	NA	2	NA	
## 5	NA	2	NA	2	NA	
## 6	NA	2	NA	2	NA	
##	B5131_W_DINING	B5132_N_DINING	B5141_W_CAR	B5142_N_CAR	B5151_W_PHONE	
## 1	2	NA	2	NA	2	
## 2	2	NA	2	NA	1	
## 3	2	NA	2	NA	1	
## 4	2	NA	2	NA	2	
## 5	2	NA	2	NA	2	
## 6	2	NA	2	NA	2	
##	B5152_N_PHONE	B5161_W_PC	B5162_N_PC	B5171_W_OVEN	B5172_N_OVEN	B5181_W_MOTOR
## 1	NA	2	NA	2	NA	2
## 2	1	2	NA	2	NA	1

## 3	1	2	NA	2	NA	1							
## 4	NA	2	NA	2	NA	2							
## 5	NA	2	NA	2	NA	2							
## 6	NA	2	NA	2	NA	2							
##	B5182_N_MOTOR	TOTEX	CTEXP	KTEXP	FOOD	CFOOD	KFOOD	CREAL	CRAL	KRAL	TRICE		
## 1	NA	53142	44759	8383	35267	33796	1471	13863	13863	0	9919		
## 2	1	189105	172913	16192	107848	107602	246	26979	26979	0	18110		
## 3	1	200101	152498	47603	107151	95206	11945	32634	24409	8225	24804		
## 4	NA	47630	31612	16018	29769	21359	8410	14612	11912	2700	13780		
## 5	NA	95456	61864	33592	42422	37557	4865	16892	15492	1400	13780		
## 6	NA	97887	56954	40933	43947	35736	8211	16770	13060	3710	13286		
##	CRICE	KRICE	TSPEC	CSPEC	KSPEC	TORDC	CORDC	KORDC	TNFAR	CNFAR	KNFAR	TOTHR	COTHR
## 1	9919	0	0	0	0	9919	9919	0	0	0	0	0	0
## 2	18110	0	0	0	0	18110	18110	0	0	0	0	0	0
## 3	16579	8225	0	0	0	24804	16579	8225	0	0	0	0	0
## 4	11080	2700	0	0	0	13780	11080	2700	0	0	0	0	0
## 5	12380	1400	0	0	0	13780	12380	1400	0	0	0	0	0
## 6	9576	3710	0	0	0	13286	9576	3710	0	0	0	0	0
##	KOTHR	TCORN	CCORN	KCORN	BREAD	CBRED	KBRED	BISCT	CBISC	KBISC	FLOUR	CFLUR	KFLUR
## 1	0	100	100	0	1300	1300	0	234	234	0	0	0	0
## 2	0	120	120	0	4160	4160	0	312	312	0	0	0	0
## 3	0	0	0	0	6240	6240	0	0	0	0	0	0	0
## 4	0	0	0	0	260	260	0	0	0	0	0	0	0
## 5	0	0	0	0	1846	1846	0	0	0	0	0	0	0
## 6	0	0	0	0	2340	2340	0	0	0	0	0	0	0
##	NCAKE	CNCKE	KNCKE	NUCLE	CNDLE	KNDLE	OCREP	COCPR	KOCPR	ROOTS	CROOT	KROOT	PTATO
## 1	520	520	0	230	230	0	1560	1560	0	0	0	0	0
## 2	0	0	0	0	0	0	4277	4277	0	580	580	0	510
## 3	810	810	0	0	0	0	780	780	0	640	640	0	540
## 4	0	0	0	0	0	0	572	572	0	115	0	115	0
## 5	0	0	0	96	96	0	1170	1170	0	120	0	120	0
## 6	0	0	0	416	416	0	728	728	0	200	125	75	125
##	CPTAT	KPTAT	CASVA	CCSVA	KCSVA	CMOTE	CCMTE	KCMTE	TGABI	CGABI	KGABI	OTRUT	COTRT
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	510	0	0	0	0	0	0	0	70	70	0	0	0
## 3	540	0	0	0	0	100	100	0	0	0	0	0	0
## 4	0	0	0	0	0	40	0	40	75	0	75	0	0
## 5	0	0	0	0	0	0	0	0	120	0	120	0	0
## 6	125	0	0	0	0	75	0	75	0	0	0	0	0
##	KOTRT	FRUIT	CFRUT	KFRUT	FRFRT	CFFRT	KFFRT	LEVEG	CLVEG	KLVEG	FRVEG	CFVEG	KFVEG
## 1	0	3880	2625	1255	544	368	176	915	619	296	1588	1074	514
## 2	0	7255	7105	150	1242	1216	26	969	949	20	3224	3157	67
## 3	0	7850	4890	2960	1350	1080	270	1400	305	1095	3050	1975	1075
## 4	0	5640	1365	4275	480	0	480	1495	15	1480	2575	650	1925
## 5	0	5265	2520	2745	1555	210	1345	625	150	475	1950	1285	665
## 6	0	4956	1490	3466	2156	150	2006	720	100	620	1550	880	670
##	BEANS	CBEAN	KBEAN	OTVEG	COVEG	KOVEG	OCROP	COCR	KOCR	FRPRE	CFRPR	KFRPR	VGP
## 1	275	186	89	229	155	74	329	223	106	0	0	0	0
## 2	587	575	12	489	479	10	744	729	15	0	0	0	0
## 3	1230	710	520	820	820	0	0	0	0	0	0	0	0
## 4	510	120	390	580	580	0	0	0	0	0	0	0	0
## 5	325	65	260	810	810	0	0	0	0	0	0	0	0
## 6	290	120	170	240	240	0	0	0	0	0	0	0	0
##	CVGPR	KVGPR	OTPRE	COTPR	KOTPR	MEAT	CMEAT	KMEAT	FCHIC	CFCHI	KFCHI	FBEEF	CFBEF

## 1	0	0	0	0	0	5140	5140	0	960	960	0	540	540
## 2	0	0	0	0	0	39024	39024	0	9360	9360	0	9360	9360
## 3	0	0	0	0	0	32480	31980	500	3620	3120	500	4940	4940
## 4	0	0	0	0	0	4025	3725	300	540	240	300	1235	1235
## 5	0	0	0	0	0	5178	4578	600	1560	960	600	0	0
## 6	0	0	0	0	0	7048	6088	960	1680	720	960	2470	2470
##	KFBEF	FPORK	CFPRK	KFPRK	OTFMT	COFMT	KOFMT	CANMT	CCNMT	KCNMT	UNCMT	CUCMT	KUCMT
## 1	0	3640	3640	0	0	0	0	0	0	0	0	0	0
## 2	0	14560	14560	0	0	0	0	534	534	0	5210	5210	0
## 3	0	14560	14560	0	0	0	0	0	0	0	9360	9360	0
## 4	0	2250	2250	0	0	0	0	0	0	0	0	0	0
## 5	0	2700	2700	0	450	450	0	468	468	0	0	0	0
## 6	0	2550	2550	0	150	150	0	198	198	0	0	0	0
##	DAIRY	CDPEG	KDPEG	TMILK	CMILK	KMILK	CONDS	CCOND	KCOND	EVAPD	CEVAP	KEVAP	POWDR
## 1	1170	1170	0	0	0	0	0	0	0	0	0	0	0
## 2	3640	3640	0	1040	1040	0	0	0	0	0	0	0	1040
## 3	3848	3848	0	1638	1638	0	0	0	0	0	0	0	1638
## 4	520	520	0	0	0	0	0	0	0	0	0	0	0
## 5	2080	2080	0	520	520	0	0	0	0	0	0	0	520
## 6	1912	1912	0	832	832	0	0	0	0	0	0	0	832
##	CPOWD	KPOWD	FRESH	CFRSH	KFRSH	ICREM	CICRM	KICRM	OTDRY	COTDR	KOTDR	TEGGS	CEGGS
## 1	0	0	0	0	0	0	0	0	0	0	0	1170	1170
## 2	1040	0	0	0	0	0	0	0	0	0	0	2600	2600
## 3	1638	0	0	0	0	0	0	0	1560	1560	0	650	650
## 4	0	0	0	0	0	0	0	0	0	0	0	520	520
## 5	520	0	0	0	0	0	0	0	0	0	0	1560	1560
## 6	832	0	0	0	0	0	0	0	780	780	0	300	300
##	KEGGS	FEGGS	CFEGG	KFEGG	PEGGS	CPEGG	KPEGG	FISHM	CFISH	KFISH	FFISH	CFFSH	KFFSH
## 1	0	1170	1170	0	0	0	0	6286	6286	0	3830	3830	0
## 2	0	2600	2600	0	0	0	0	8990	8990	0	7910	7910	0
## 3	0	650	650	0	0	0	0	16840	16840	0	13050	13050	0
## 4	0	520	520	0	0	0	0	2842	1952	890	2080	1190	890
## 5	0	1560	1560	0	0	0	0	6340	6340	0	4790	4790	0
## 6	0	300	300	0	0	0	0	6161	6161	0	4023	4023	0
##	CNFSH	CCFSH	KCFSH	DRFSH	CDFSH	KDFSH	SLFSH	CSFSH	KSFSH	OTMRN	COMRN	KOMRN	COFCT
## 1	1430	1430	0	350	350	0	676	676	0	0	0	0	956
## 2	132	132	0	350	350	0	598	598	0	0	0	0	3230
## 3	0	0	0	2840	2840	0	950	950	0	0	0	0	991
## 4	312	312	0	0	0	0	450	450	0	0	0	0	240
## 5	0	0	0	860	860	0	690	690	0	0	0	0	1094
## 6	624	624	0	780	780	0	734	734	0	0	0	0	2524
##	CCOFT	KCOFT	COFEE	CCOFE	KCOFE	COFPR	CCFPR	KCFPR	COFBN	CCFBN	KCFBN	COCOA	CCCOA
## 1	884	72	956	884	72	956	884	72	0	0	0	0	0
## 2	3198	32	1786	1768	18	1786	1768	18	0	0	0	1444	1430
## 3	991	0	991	991	0	991	991	0	0	0	0	0	0
## 4	240	0	240	240	0	240	240	0	0	0	0	0	0
## 5	1094	0	1094	1094	0	1094	1094	0	0	0	0	0	0
## 6	2524	0	1094	1094	0	1094	1094	0	0	0	0	1430	1430
##	KCCOA	COAPT	CCOFT	KCOFT	COAPR	CCOPR	KCOPR	COABN	CCOBN	KCOBN	TEA	CTEA	KTEA
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	14	1444	1430	14	0	0	0	0	0	0	0	0	0
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	0	0	0	0	0	0	0	0	0	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	0

## 6	0	1430	1430	0	0	0	0	0	0	0	0	0	0
##	TEAPR	CTEPR	KTEPR	TEALV	CTELV	KTELV	NONAL	CNALC	KNALC	CARBD	CCARB	KCARB	NCARB
## 1	0	0	0	0	0	0	528	528	0	528	528	0	0
## 2	0	0	0	0	0	0	2466	2466	0	2340	2340	0	126
## 3	0	0	0	0	0	0	7462	7462	0	7462	7462	0	0
## 4	0	0	0	0	0	0	168	168	0	168	168	0	0
## 5	0	0	0	0	0	0	2052	2052	0	2052	2052	0	0
## 6	0	0	0	0	0	0	1650	1650	0	1650	1650	0	0
##	CNCAR	KNCAR	OTHDR	COTHD	KOTHD	BOTLE	CBOTLE	KBOTLE	FDNEC	CFNEC	KFNEC	SUGAR	
## 1	0	0	0	0	0	0	0	0	3444	3300	144	1062	
## 2	126	0	0	0	0	0	0	0	6006	5942	64	1045	
## 3	0	0	0	0	0	0	0	0	4406	4146	260	1300	
## 4	0	0	0	0	0	0	0	0	1607	1477	130	120	
## 5	0	0	0	0	0	0	0	0	3401	3401	0	1033	
## 6	0	0	0	0	0	0	0	0	2726	2726	0	608	
##	CSGAR	KSGAR	SUGPR	CSUPR	KSUPR	CKOIL	CCOIL	KCOIL	MARGN	CMARG	KMARG	SAUCE	CSAUC
## 1	1040	22	0	0	0	956	936	20	0	0	0	123	120
## 2	1040	5	0	0	0	712	708	4	0	0	0	780	776
## 3	1300	0	0	0	0	936	936	0	0	0	0	0	0
## 4	120	0	0	0	0	474	474	0	0	0	0	10	10
## 5	1033	0	0	0	0	988	988	0	0	0	0	24	24
## 6	608	0	0	0	0	588	588	0	0	0	0	0	0
##	KSAUC	TSALT	CSALT	KSALT	OTSPC	COSPC	KOSPC	MLOUT	CMOUT	KMOUT	OFNEC	COFNC	KOFNC
## 1	3	112	110	2	494	484	10	0	0	0	697	610	87
## 2	4	191	190	1	1102	1096	6	0	0	0	2176	2132	44
## 3	0	160	160	0	2010	1750	260	0	0	0	0	0	0
## 4	0	112	112	0	891	761	130	0	0	0	0	0	0
## 5	0	69	69	0	1287	1287	0	0	0	0	0	0	0
## 6	0	128	128	0	1402	1402	0	0	0	0	0	0	0
##	FHOME	CFHME	KFHME	FDOUT	CFOUT	KFOUT	MLSCH	CMLSC	KMLSC	MLWRK	CMLWR	KMLWR	
## 1	35267	33796	1471	0	0	0	0	0	0	0	0	0	0
## 2	98170	97924	246	9678	9678	0	0	0	0	0	0	0	0
## 3	107151	95206	11945	0	0	0	0	0	0	0	0	0	0
## 4	29769	21359	8410	0	0	0	0	0	0	0	0	0	0
## 5	42422	37557	4865	0	0	0	0	0	0	0	0	0	0
## 6	43947	35736	8211	0	0	0	0	0	0	0	0	0	0
##	MLRES	CMLRE	KMLRE	SNACK	CSNAK	KSNAK	ALBEV	CALCB	KALCB	TBEER	CBEER	KBEER	NWINE
## 1	0	0	0	0	0	0	1372	1300	72	1372	1300	72	0
## 2	0	0	0	9678	9678	0	32	0	32	32	0	32	0
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	0	0	0	986	986	0	0	0	0	0
## 5	0	0	0	0	0	0	728	728	0	0	0	0	0
## 6	0	0	0	0	0	0	702	702	0	0	0	0	0
##	CNWNE	KNWNE	OTBEV	COTBV	KOTBV	NFOOD	CNFOOD	KNFOOD	TBCCO	CTOBC	KTOBC	CIGRT	
## 1	0	0	0	0	0	17875	10963	6912	384	312	72	384	
## 2	0	0	0	0	0	81257	65311	15946	32	0	32	32	
## 3	0	0	0	0	0	92950	57292	35658	0	0	0	0	
## 4	0	0	986	986	0	17861	10253	7608	3094	3094	0	3094	
## 5	0	0	728	728	0	53034	24307	28727	4004	4004	0	4004	
## 6	0	0	702	702	0	53940	21218	32722	260	260	0	0	
##	CCGRT	KCGRT	CIGAR	CCGAR	KCGAR	OTTOB	COTOB	KOTOB	FUEL1	CFUEL	KFUEL	A1022	A1023
## 1	312	72	0	0	0	0	0	0	3474	876	2598	0	0
## 2	0	32	0	0	0	0	0	0	20598	20100	498	2922	2922
## 3	0	0	0	0	0	0	0	0	15990	15990	0	540	540

## 4	3094	0	0	0	0	0	0	0	2892	420	2472	0	0
## 5	4004	0	0	0	0	0	0	0	7914	5742	2172	162	162
## 6	0	0	0	0	0	260	260	0	14628	12606	2022	0	0
##	A1024	A1032	A1033	A1034	A1042	A1043	A1044	A1052	A1053	A1054	A1062	A1063	A1064
## 1	0	2598	0	2598	0	0	0	876	876	0	0	0	0
## 2	0	498	0	498	1854	1854	0	0	0	0	11634	11634	0
## 3	0	0	0	0	4650	4650	0	0	0	0	6000	6000	0
## 4	0	2472	0	2472	0	0	0	420	420	0	0	0	0
## 5	0	2172	0	2172	720	720	0	36	36	0	2976	2976	0
## 6	0	2022	0	2022	1350	1350	0	36	36	0	6990	6990	0
##	A1072	A1073	A1074	A1082	A1083	A1084	A1092	A1093	A1094	TRCOM	CTRCM	KTRCM	A2022
## 1	0	0	0	0	0	0	0	0	0	684	684	0	684
## 2	0	0	0	3690	3690	0	0	0	0	14094	14094	0	7164
## 3	0	0	0	4800	4800	0	0	0	0	7206	6918	288	0
## 4	0	0	0	0	0	0	0	0	0	264	120	144	120
## 5	0	0	0	1848	1848	0	0	0	0	4362	4362	0	4362
## 6	0	0	0	4230	4230	0	0	0	0	720	720	0	0
##	A2032	A2042	A2052	A2062	A2072	A2082	A2092	A2102	A2112	A2122	A2132	A3022	A3032
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	0	0	3420	0	0	0	3510	0	0	0	0	0	0
## 3	0	0	2598	360	0	0	3960	0	0	0	0	288	0
## 4	0	0	0	0	0	0	0	0	0	0	0	144	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	0
## 6	0	0	720	0	0	0	0	0	0	0	0	0	0
##	A3042	A3052	A3062	A3072	A3082	A3092	A3102	A3112	A3122	A3132	HOPER	CHHOP	KHHOP
## 1	0	0	0	0	0	0	0	0	0	0	1128	1128	0
## 2	0	0	0	0	0	0	0	0	0	0	2430	2430	0
## 3	0	0	0	0	0	0	0	0	0	0	3534	3534	0
## 4	0	0	0	0	0	0	0	0	0	0	894	894	0
## 5	0	0	0	0	0	0	0	0	0	0	870	870	0
## 6	0	0	0	0	0	0	0	0	0	0	1236	1236	0
##	A4022	A4023	A4024	A4032	A4033	A4034	A4042	A4043	A4044	A4052	A4053	A4054	A4062
## 1	876	876	0	0	0	0	0	0	0	0	0	0	0
## 2	1686	1686	0	0	0	0	54	54	0	0	0	0	270
## 3	2718	2718	0	0	0	0	0	0	0	0	0	0	768
## 4	774	774	0	0	0	0	0	0	0	0	0	0	0
## 5	762	762	0	0	0	0	0	0	0	0	0	0	0
## 6	1152	1152	0	0	0	0	0	0	0	0	0	0	0
##	A4063	A4064	A4072	A4073	A4074	A4082	A4083	A4084	A4092	A4093	A4094	A4112	A4113
## 1	0	0	0	0	0	0	0	0	252	252	0	0	0
## 2	270	0	0	0	0	0	0	0	420	420	0	0	0
## 3	768	0	0	0	0	0	0	0	48	48	0	0	0
## 4	0	0	0	0	0	0	0	0	120	120	0	0	0
## 5	0	0	0	0	0	0	0	0	108	108	0	0	0
## 6	0	0	0	0	0	0	0	0	84	84	0	0	0
##	A4114	A4122	A4123	A4124	DSERV	CDSER	KDSER	A4132	A4133	A4134	A4142	A4143	A4144
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	0	0	0	0	0	0	0	0	0	0	0	0	0
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	0	0	0	0	0	0	0	0	0	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	0
## 6	0	0	0	0	0	0	0	0	0	0	0	0	0
##	A4152	A4153	A4154	A4162	A4163	A4164	PRCRE	CPCRE	KPCRE	A5022	A5082	A5032	A5092
## 1	0	0	0	0	0	0	3066	2496	570	2496	0	0	0

## 2	0	0	0	0	0	0	9672	6288	3384	5988	2604	0	0
## 3	0	0	0	0	0	0	14976	9606	5370	9006	3990	0	0
## 4	0	0	0	0	0	0	2568	1176	1392	1176	192	0	0
## 5	0	0	0	0	0	0	6030	3630	2400	3630	0	0	0
## 6	0	0	0	0	0	0	2490	1890	600	1788	0	0	0
##	A5042	A5102	A5052	A5112	A5062	A5122	BCARE	GBCARE	ACARE	GACARE	CLOTH	CCLOT	
## 1	0	102	0	468	0	0	0	0	0	0	554	554	
## 2	300	0	0	780	0	0	0	0	0	0	4675	4675	
## 3	300	1080	300	300	0	0	0	0	0	0	0	0	
## 4	0	0	0	1200	0	0	0	0	0	0	330	330	
## 5	0	1800	0	600	0	0	0	0	0	0	365	210	
## 6	0	0	102	600	0	0	0	0	0	0	850	850	
##	KCLOT	A6022	A6032	A6042	A6052	A6062	A6072	A6082	A6092	A6102	A6112	A6122	A6132
## 1	0	0	0	0	0	0	310	115	129	0	0	0	0
## 2	0	1020	2230	0	0	0	415	1010	0	0	0	0	0
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	100	0	0	0	0	80	150	0	0	0
## 5	155	0	0	0	0	0	0	210	0	0	0	0	0
## 6	0	0	0	0	0	850	0	0	0	0	0	0	0
##	A7022	A7032	A7042	A7052	A7062	A7072	A7082	A7092	A7102	A7112	A7122	A7132	EDUC
## 1	0	0	0	0	0	0	0	0	0	0	0	0	644
## 2	0	0	0	0	0	0	0	0	0	0	0	0	7565
## 3	0	0	0	0	0	0	0	0	0	0	0	0	16130
## 4	0	0	0	0	0	0	0	0	0	0	0	0	111
## 5	0	0	0	0	0	15	140	0	0	0	0	0	0
## 6	0	0	0	0	0	0	0	0	0	0	0	0	0
##	CEDUC	KEDUC	A8022	A8032	A8042	A8052	A8062	A8082	A8092	A8102	A8112	A8122	RCRTN
## 1	644	0	160	0	0	484	0	0	0	0	0	0	0
## 2	7565	0	4020	0	2800	745	0	0	0	0	0	0	0
## 3	16130	0	15460	0	0	670	0	0	0	0	0	0	0
## 4	111	0	0	0	0	111	0	0	0	0	0	0	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	0
## 6	0	0	0	0	0	0	0	0	0	0	0	0	0
##	CRCTN	KRCTN	A9022	A9032	A9042	A9052	A9062	A9082	A9092	A9102	A9112	A9122	MEDIC
## 1	0	0	0	0	0	0	0	0	0	0	0	0	371
## 2	0	0	0	0	0	0	0	0	0	0	0	0	3211
## 3	0	0	0	0	0	0	0	0	0	0	0	0	1820
## 4	0	0	0	0	0	0	0	0	0	0	0	0	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	672
## 6	0	0	0	0	0	0	0	0	0	0	0	0	160
##	CMEDI	KMEDI	B0022	B0032	B0042	B0052	B0062	B0072	PILLS	FSUPP	B0092	B0102	B0112
## 1	371	0	371	0	0	0	0	0	0	0	0	0	0
## 2	3211	0	335	2500	300	0	76	0	0	0	0	0	0
## 3	1820	0	1820	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	0	0	0	0	0	0	0	0	0	0
## 5	672	0	672	0	0	0	0	0	0	0	0	0	0
## 6	60	100	60	0	0	0	0	0	0	0	100	0	0
##	B0122	B0132	B0142	GPILL	GFSUP	NDFUR	CNDUR	KNDUR	B1022	B1032	B1042	B1062	B1072
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	0	0	0	0	0	0	0	0	0	0	0	0	0
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	0	0	0	0	0	0	0	0	0	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	0
## 6	0	0	0	0	0	0	0	0	0	0	0	0	0

##	B1082	DUFUR	CDFUR	KDFUR	B2022	B2102	B3022	B2032	B2112	B3032	B2042	B2122	B3042
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	0	0	0	0	0	0	0	0	0	0	0	0	0
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	0	0	0	0	0	0	0	0	0	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	0
## 6	0	0	0	0	0	0	0	0	0	0	0	0	0
##	B2052	B2132	B3052	B2062	B2142	B3062	B2072	B2152	B3072	B2082	B2162	B3082	TAXES
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	0	0	0	0	0	0	0	0	0	0	0	0	1200
## 3	0	0	0	0	0	0	0	0	0	0	0	0	1219
## 4	0	0	0	0	0	0	0	0	0	0	0	0	0
## 5	0	0	0	0	0	0	0	0	0	0	0	0	717
## 6	0	0	0	0	0	0	0	0	0	0	0	0	722
##	B3102	B3112	B3122	B3132	HOUSE	ACRNT	RNTVL	RNTVX	RPAIR	CREPR	KREPR	B6022	B6102
## 1	0	0	0	0	3600	0	3600	0	0	0	0	0	0
## 2	0	0	1200	0	12000	0	12000	0	0	0	0	0	0
## 3	284	925	0	10	30000	0	30000	0	0	0	0	0	0
## 4	0	0	0	0	3600	0	3600	0	0	0	0	0	0
## 5	0	712	0	5	24000	0	24000	0	0	0	0	0	0
## 6	0	712	0	10	30000	0	30000	0	0	0	0	0	0
##	B6032	B6112	B6042	B6122	B6052	B6132	B6062	B6142	B6072	B6152	B6082	B6162	OCCSN
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	0	0	0	0	0	0	0	0	0	0	0	0	4500
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	0	0	0	0	0	0	0	0	0	0	0	0	650
## 5	0	0	0	0	0	0	0	0	0	0	0	0	1200
## 6	0	0	0	0	0	0	0	0	0	0	0	0	150
##	COCCN	KOCCN	B7022	B7092	B7032	B7102	B7042	B7112	B7052	B7122	B7062	B7132	B7072
## 1	0	0	0	0	0	0	0	0	0	0	0	0	0
## 2	4500	0	4500	0	0	0	0	0	0	0	0	0	0
## 3	0	0	0	0	0	0	0	0	0	0	0	0	0
## 4	650	0	500	0	150	0	0	0	0	0	0	0	0
## 5	1200	0	1000	0	200	0	0	0	0	0	0	0	0
## 6	150	0	150	0	0	0	0	0	0	0	0	0	0
##	B7142	GFTOT	CGIFT	KGIFT	B8022	B8023	B8032	B8033	B8042	B8043	B8052	B8053	OTHEX
## 1	0	0	0	0	0	0	0	0	0	0	0	0	2598
## 2	0	750	750	0	500	0	250	0	0	0	0	0	498
## 3	0	0	0	0	0	0	0	0	0	0	0	0	2075
## 4	0	0	0	0	0	0	0	0	0	0	0	0	2472
## 5	0	0	0	0	0	0	0	0	0	0	0	0	2172
## 6	0	0	0	0	0	0	0	0	0	0	0	0	2022
##	B8072	B8082	B8092	OTDIS	B9022	B9032	B9042	B9052	B9062	B9072	B9092	B9102	B9082
## 1	0	0	2598	0	0	0	0	0	0	0	0	0	0
## 2	0	0	498	0	0	0	0	0	0	0	0	0	0
## 3	1700	0	375	142	0	0	0	0	0	0	0	0	142
## 4	0	0	2472	0	0	0	0	0	0	0	0	0	0
## 5	0	0	2172	0	0	0	0	0	0	0	0	0	0
## 6	0	0	2022	0	0	0	0	0	0	0	0	0	0
##	TOTDI	CASDI	KINDI	TOREC		TOINC	WAGES	WSAG		WSNAG	OTHIN	NETSH	
## 1	53142	44759	8383	51685		51623.33	40812.67	0		40812.67	10685		0
## 2	189105	172913	16192	190156		172079.33	0.00	0		0.00	145434		15000
## 3	200243	152640	47603	216003		196118.00	96044.00	0		96044.00	82878		31960
## 4	47630	31612	16018	47936		30555.33	3080.00	3080		0.00	10796		0

```

## 5  95456  61864 33592 142125 133475.00 76638.67    0 76638.67  46805    0
## 6  97887  56954 40933  90467  72495.33    0.00    0    0.00 55631 15231
##    CONAB CONDO RENTL INTRS PNSNS DVDND IFAMS REGFT OSINC EAINC EACFGGRS
## 1      0  2000    0    0    0    0  3480 1605    0    0    0
## 2 114000    0    0    0    0    0  1050 3384    0 26612  19860
## 3  15000    0    0    0    0    0    0  5918    0 18480    0
## 4      0    0    0    0    0    0  5290 1906    0 16940  19100
## 5   8000  7200    0    0    0    0  5050 2555    0 10340  14890
## 6   5000    0    0    0    0    0  4700  700    0 16860  25160
##    EALPRGRS EAFISGRS EAFORGRS EATRDGRS EAMFGGRS EACPSGRS EATCSGRS EAMNGGRS
## 1      0      0      0      0      0      0      0      0
## 2      0      0      0      0      12000  13002      0      0
## 3      0      0      0      0      0      0      62400      0
## 4      0      0      0      0      0      0      0      0
## 5      0      0      0      0      0      0      0      0
## 6      0      0      0      0      0      0      0      0
##    EACONGRS EANEGRS EACFGEXP EALPREXP EAFISEXP EAFOREXP EATRDEXP EAMFGEXP
## 1      0      0      0      0      0      0      0      0
## 2      0      0      6250      0      0      0      0      6000
## 3      0      0      0      0      0      0      0      0
## 4      0      0      2160      0      0      0      0      0
## 5      0      0      4550      0      0      0      0      0
## 6      0      0      8300      0      0      0      0      0
##    EACPSEXP EATCSEXP EAMNGEXP EACONEXP EANECEXP EACFG EALPR EAFIS EAFOR EATRD
## 1      0      0      0      0      0      0      0      0      0      0
## 2      6000      0      0      0      0      13610      0      0      0      0
## 3      0      43920      0      0      0      0      0      0      0      0
## 4      0      0      0      0      0      16940      0      0      0      0
## 5      0      0      0      0      0      10340      0      0      0      0
## 6      0      0      0      0      0      16860      0      0      0      0
##    EAMFG EACPS EATCS EAMNG EACON EANECE EA_LOSS OTREC ILOAN WDRAW WINNG PRFIT
## 1      0      0      0      0      0      0      0      0      0      0      0
## 2  6000  7002      0      0      0      0      18110      0 18110      0      0
## 3      0      0 18480      0      0      0      18813      0 18813      0      0
## 4      0      0      0      0      0      0      17080      0 11080  6000      0
## 5      0      0      0      0      0      0      8580      0      0  8580      0
## 6      0      0      0      0      0      0      17976      0 17976      0      0
##    BKPAY INHRT OTHRE AGINC  NAGIN NATDC REGDC NATPC REGPC MAJSR MINSR AGIND
## 1      0      0      0  3178  48507      2      2      1      1      1      2      2
## 2      0      0      0 14358 157688      7      7      7      6      3     15      2
## 3      0      0      0      0 197190      8      7      5      5      1      2      2
## 4      0      0      0 23722   7134      1      1      1      1      2      3      1
## 5      0      0      0 13262 120283      6      6      6      5      1      2      2
## 6      0      0      0 19842  52649      4      3      4      4      3     21      2

```

```

# Optionally, export the anonymized data to a CSV file
write.csv(final_data, "anonymized_FIES2006.csv", row.names = FALSE)

```

COMMENT:

this process ensures that the dataset is anonymized and ready for use while protecting individual privacy. The export to CSV provides a convenient way to share or store the anonymized data for future reference.