Comparison of DemoDiff Releases
2.0 and 3.0

Rainer Walke (walke@demogr.mpg.de)

For additional material see www.demogr.mpg.de/tr/
Comparison of DemoDiff Releases 2.0 and 3.0

Rainer Walke, MPIDF Rostock*

2013-Aug-26

Abstract

In this Technical Report, results of the program compareFinRaw are presented. compareFinRaw is a tool that is particularly useful for comparing large data sets. It is typically used to compare different releases of the same data. In this example, we compare two different releases of the project DemoDiff. To be more precise, we compare the five revised files of DemoDiff Release 3.0 with the older Release 2.0 files. It should provide users of the DemoDiff data a comprehensive overview on all manipulations of the data that have occurred from Release 2.0 to Release 3.0.

Keywords    data analysis, data comparability, data evaluation, data processing, software

1. Introduction

In a Technical Report by Walke and Müller [TR-2012-003] we have described a procedure to compare two datasets with little conditions. Only the ID had to be the same in each row of both datasets. The program that compares the datasets had been named compareFinRaw. It compares each column (variable) from one set with each column from the second data set. It furthermore checks whether there are bijective mappings between variables of the two data sets. If there is no direct mapping it computes how much the levels of the variables have to be changed to get a bijective mapping.

In this report, we use compareFinRaw to compare two different releases of the project DemoDiff. We are comparing Release 2.0 [DemoDiff 2.0] and Release 3.0 [DemoDiff 3.0]. Both are available by GESIS (www.gesis.org). We compare the following data sets: anchor1_DD.dta, anchor2_DD.dta, partner1_DD.dta, partner2_DD.dta and weights.dta1.

We are using the statistical package R [R 2.15] for all the computations and RStudio, knitR and MiKTeX for the documentation. The output of these programs is a readable list that reports the number and the names of variables that completely match between the two releases, i.e. have the same name and same content. The program also identifies all modified variables such as renamed variables, bijective recoded variables and variables with marked differences. Furthermore we get the names of all variables without variation at all. We hope, that this material helps users of DemoDiff to get a clear picture on what has been changed between the two releases. It should be noted that, depending on the number of variables compareFinRaw takes minutes or hours to compare the files.  

---

*MPIDF, Konrad-Zuse-Straße 1, D-18057 Rostock, Germany. E-Mail: walke@demogr.mpg.de

1There are further files in the releases. However, they have been omitted because they are not suitable for comparison.

2In our case, it took 7 hours to run the program for the largest data set (anchor2_DD.dta)
2. Classification

The idea is to classify all variables in data set A in respect to their relations to variables from data set B.

A typical solution would be first to check whether a variable with the same name is available in data set B. If this is true we could distinguish between identical and non identical content. This way we would classify all variable from set A into one of three possible categories.

This would give only a rough picture of the differences between the data sets. To detect renamed and recoded variables as well we use a more sophisticated classification scheme.

classification tree

1. The selected variable from A does not have variation.
   1.1 [ x ] No variable with the same name is available in data set B.
   1.2 One variable with the same name is available in data set B.
      1.21 [ n ] Both variables are not identical.
      1.22 [ ni ] Both variables are identical.

2. The selected variable from A does have variation.
   2.1 There exists no bivariate mapping to one of the variables in data set B.
      2.11 [ v ] No variable with the same name is available in data set B.
      2.12 [ vn ] One variable with the same name is available in data set B.
   2.2 There exists at least one bivariate mapping to one of the variables in data set B.
      2.21 There are more than one bivariate mappings.
         2.211 All mapped variables from B do have a different name.
            2.2111 [ vb ] No mapped variable from B is identical.
            2.2112 [ vbi ] At least one mapped variable from B is identical.
            2.212 One mapped variable from B has the same name.
               2.2121 [ vb1 ] Both variables are not identical.
               2.2122 [ vb1i ] Both variables are identical.
      2.22 There exists exactly one bivariate mapping.
         2.221 The mapped variable from B has a different name.
            2.2211 [ vb1 ] Both variables are not identical.
            2.2212 [ vb1i ] Both variables are identical.
         2.222 The mapped variable from B has the same name.
            2.2221 [ vb1n ] Both variables are not identical.
            2.2222 [ vb1ni ] Both variables are identical.

Every variable out of A will be assign to exactly one of these 13 classes.
3. Results and Summary

The appendix contains the documented code. For each of the five data sets the code follows the same logic.

1. Analyze all variables without variation
2. Analyze all variables with variation, but without a bijective mapping
3. Analyze all variables with variation and with a bijective mapping
4. Provide a comparison summary

As an example we display here the summary output that our program has produced for the data set anchor2_DD.dta.

<table>
<thead>
<tr>
<th>class</th>
<th>Release 2</th>
<th>Release 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>ni</td>
<td>612</td>
<td>612</td>
</tr>
<tr>
<td>v</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>vn</td>
<td>369</td>
<td>359</td>
</tr>
<tr>
<td>vb</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>vbi</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>vbn</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>vbni</td>
<td>954</td>
<td>956</td>
</tr>
<tr>
<td>vb1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>vb1i</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>vb1n</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>vb1ni</td>
<td>1262</td>
<td>1260</td>
</tr>
<tr>
<td>sum</td>
<td>3246</td>
<td>3244</td>
</tr>
</tbody>
</table>

The results may be interpreted as follows:
Most variables are unchanged (ni, vbni, vb1ni). 612 variables do not have variation and are identical. 954 + 1262 variables (vbni, vb1ni, Release 2) are unchanged, but have variation. There is a weak probability that 954 variables (vbni) show some redundancy.

8 constant variables (x, Release 2) have been dropped. One variable (v, Release 2) with variation has been dropped as well.

15 variables (n, Release 2) have been gained variation or have changed to another constant value.

2 variables (vb, Release 2) have been (less probably) renamed and recoded. 11 variables (vbi, Release 2) have been (less probably) renamed. 2 variables (vbn, Release 2) have been (probably) recoded.

4 variables (vb1, Release 2) have been (probably) renamed and recoded. 3 variables (vb1i, Release 2) have been (probably) renamed. 3 variables (vb1n, Release 2) have been (probably) recoded.

369 variables (vn, Release 2) had been changed between releases. They have the same name, but there exist no bijective mapping. The provided Levenshtein distance (appendix) gives a rough idea whether only some cases have been changed or not.

4. Acknowledgements

The report was supported by the Max Planck Institute for Demographic Research, Rostock, Germany. The author would like to thank Michaela Kreyenfeld, Rüdiger Lenke and Tom Hensel for valuable comments on the paper. I thank Andreas Müller for helping with the compareFinRaw.r code preparation.
## A. anchor1

Here we compare the data set anchor1_dd.dta from DemoDiff Release 2.0 with the same data set from DemoDiff Release 3.0.

We have to load the produced data collection. Release 2.0 files are denoted with R and Raw, release 3.0 files with F and Fin, respectively.

```r
rm(list = ls())
duplicStrict <- function(A) {
  return(duplicated(A) | duplicated(A, fromLast = TRUE))
}
load("..\compareR2R3\anchor1\Results\compareFinRaw.RData")
```

### A.1. no variation

We compare the variables without variation first.

```r
(R.x <- sum(!(RawNames[Rnv] %in% FinNames)))
## [1] 1

(F.x <- sum(!(FinNames[Fnv] %in% RawNames)))
## [1] 0

# added or lost variance
(R.nA <- sum(!(RawNames[Rnv] %in% FinNames[Fnv])) - R.x)
## [1] 0

(F.nA <- sum(!(FinNames[Fnv] %in% RawNames[Rnv])) - F.x)
## [1] 0

Rnv2 <- data.frame(Rnv.tab)
Fnv2 <- data.frame(Fnv.tab)
Rnv2$nm <- gsub("R$", "", row.names(Rnv2))
Fnv2$nm <- gsub("F$", "", row.names(Fnv2))
RFnv2 <- merge(Rnv2, Fnv2, by = c("nm"), all = TRUE)

(R.ni <- sum((RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 423

F.ni <- R.ni

# changed values
(R.nB <- sum(!(RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
```
Release 2.0  For 1 variables without variation (k10age) exist no variables with the same name in the comparison data set. 423 variables without variation are identical in both data sets.

Release 3.0  423 variables without variation are identical in both data sets.

A.2. variation, but no bijective mapping

rb <- results.bijec
RL.bi <- RawNames[!Rnv] %in% gsub("R$", ",", as.character(rb$Raw.c.nm))
FL.bi <- FinNames[!Fnv] %in% gsub("F$", ",", as.character(rb$Fin.c.nm))
sum(!RL.bi)
## [1] 43

sum(!FL.bi)
## [1] 48
(R.v <- sum(!(RawNames[!Rnv][!RL.bi] %in% FinNames)))
## [1] 0

RawNames[!Rnv][!RL.bi][!(RawNames[!Rnv][!RL.bi] %in% FinNames)]

## character(0)

(R.vn <- sum(RawNames[!Rnv][!RL.bi] %in% FinNames))
## [1] 43

RawNames[!Rnv][!RL.bi][RawNames[!Rnv][!RL.bi] %in% FinNames]

## [1] "sex"    "dobm"   "sd4g"    "sdplm"    "sdply"
## [6] "frt11v1i1" "frt11v1i2" "frt11v1i3" "frt11v1i4" "frt11v1i5"
## [11] "frt11v1i6" "frt11v1i7" "frt11v1i8" "frt11v2i1" "frt11v2i2"
## [16] "frt11v2i3" "frt11v2i4" "frt11v2i5" "frt11v2i6" "frt11v2i7"
## [21] "frt11v2i8" "flag1"   "flag18"   "sex_gen"   "psex_gen"
## [26] "dobm_gen" "pdoby_gen" "pdobm_gen" "age"     "page"
## [31] "mage"    "fage"    "k1age"    "k2age"    "homosex"
## [36] "infertile" "pregnant" "hhsizemrd" "othmrd"   "hhcomp"
## [41] "isei"    "siops"   "hhincgcee"

(F.v <- sum(!(FinNames[!Fnv][!FL.bi] %in% RawNames)))
## [1] 5

FinNames[!Fnv][!FL.bi][!(FinNames[!Fnv][!FL.bi] %in% RawNames)]

## [1] "ykage"   "ykid"   "flag_isco08_kldb2010"
## [4] "kldb2010" "isco08"

(F.vn <- sum(FinNames[!Fnv][!FL.bi] %in% RawNames))
## [1] 43

FinNames[!Fnv][!FL.bi][FinNames[!Fnv][!FL.bi] %in% RawNames]

## [1] "sex"    "dobm"   "sd4g"    "sdplm"    "sdply"
## [6] "frt11v1i1" "frt11v1i2" "frt11v1i3" "frt11v1i4" "frt11v1i5"
## [11] "frt11v1i6" "frt11v1i7" "frt11v1i8" "frt11v2i1" "frt11v2i2"
## [16] "frt11v2i3" "frt11v2i4" "frt11v2i5" "frt11v2i6" "frt11v2i7"
## [21] "frt11v2i8" "flag1"   "flag18"   "sex_gen"   "psex_gen"
## [26] "dobm_gen" "pdoby_gen" "pdobm_gen" "age"     "page"
## [31] "mage"    "fage"    "k1age"    "k2age"    "homosex"
## [36] "infertile" "pregnant" "hhsizemrd" "othmrd"   "hhcomp"
## [41] "isei"    "siops"   "hhincgcee"
Release 2.0  43 variables share the name with one of the variables in Release 3.0 at least.

Release 3.0  For 5 variables with variation but without any bijective mapping exist no variables with the same name in the Release 2.0 data. 43 variables share the name with one of the variables in Release 2.0 at least.

We compare all variables pairs (with variation) which share the same name but are not connected with a bijective mapping. Those variables imply different information. DemoDiff data user should check whether they are using those variables for possible effects on their research. (The number of pairs may differ from 43 and 43).

```r
results$same.name <-
gsub("R$","",as.character(results$Raw.c.nm)) == gsub("F$","",as.character(results$Fin.c.nm))
sum(results$same.name & results$map.di>0)
## [1] 43

print(results[(results$same.name & results$map.di>0),c(2,3,5,6,7,8)],row.names=FALSE)
##    Fin.c.nm Fin.c.ls Raw.c.nm Raw.c.ls map.di ed.di.sum
##  sexF  2   sexR  2      4      4      2
##  dobmF 12  dobmR 12      10     10      7
##  sd4gF  3    sd4gR  3      4      4      4
##  sdplmF 17   sdplmR 17      20     20      4
##  sdplyF 45   sdplyR 45      30     30      25
## frt11v1i1F 10  frt11v1i1R  9    1    1
## frt11v1i2F 10  frt11v1i2R  9    1    1
## frt11v1i3F 10  frt11v1i3R  9    1    1
## frt11v1i4F 10  frt11v1i4R  9    1    1
## frt11v1i5F 10  frt11v1i5R  9    1    1
## frt11v1i6F 10  frt11v1i6R  9    1    1
## frt11v1i7F  9   frt11v1i7R  8    1    1
## frt11v1i8F 10  frt11v1i8R  9    1    1
## frt11v2i1F  6   frt11v2i1R  5    1    1
## frt11v2i2F  6   frt11v2i2R  5    1    1
## frt11v2i3F  6   frt11v2i3R  5    1    1
## frt11v2i4F  6   frt11v2i4R  5    1    1
## frt11v2i5F  6   frt11v2i5R  5    1    1
## frt11v2i6F  6   frt11v2i6R  5    1    1
## frt11v2i7F  6   frt11v2i7R  5    1    1
## frt11v2i8F  6   frt11v2i8R  5    1    1
##  flag1F  2    flag1R  2       4    4
## flag18F  3   flag18R  3       2    2
## sex_genF  2   sex_genR  2       4    4
## psex_genF  3   psex_genR  3       4    4
## dobm_genF 12  dobm_genR 12      10     10      7
## pdoby_genF 45  pdoby_genR 45      30     30      25
## pdobm_genF 17  pdobm_genR 17      20     20      17
##  ageF   9    ageR   9       2    2
##  pageF  45   pageR  45       36    36      25
##  mageF  43   mageR  43       16    16      13
##  fageF  46   fageR  46       20    20      12
## k1ageF  29   k1ageR  29       4    4
## k2ageF  26   k2ageR  26       4    4
## homosexF 3   homosexR  3       4    4
```

## infertileF 3 infertileR 3 2 2
## pregnantF 4 pregnantR 4 4 6
## hhsizemrdF 9 hhsizemrdR 9 2 1
## othmrdF 6 othmrdR 6 2 1
## hhcompF 15 hhcompR 15 2 1
## iseiF 174 iseiR 60 462 2486
## siopsF 177 siopsR 57 478 2529
## hhincgceeF 330 hhincgceeR 331 1 7

### A.3. variation and bijective mapping

```r
# mark all identical cases
rb$ident <- (rb$ed.di.sum == 0)
# mark all not duplicated raw variables
rb$dup.Raw <- !duplicStrict(rb$Raw.c.nm)
# mark all not duplicated fin variables
rb$dup.Fin <- !duplicStrict(rb$Fin.c.nm)
# mark all equal name pairs
rb$same.name <- gsub("$", "", as.character(rb$Raw.c.nm)) == gsub("$", "", as.character(rb$Fin.c.nm))

rb$Raw.same.name <- (ave(rb$same.name, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.same.name <- (ave(rb$same.name, rb$Fin.c.nm, FUN = sum) > 0)

rb$Raw.ident <- (ave(rb$ident, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.ident <- (ave(rb$ident, rb$Fin.c.nm, FUN = sum) > 0)

(R.vb <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"])))
## [1] 0

(R.vbi <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"])))
## [1] 0

(R.vbn <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"])))
## [1] 0

(R.vbni <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"])))
## [1] 225

R.vb1 <- length(unique(rb[rb$dup.Raw & !rb$same.name & !rb$ident, "Raw.c.nm"]))
as.character(unique(rb[rb$dup.Raw & !rb$same.name & !rb$ident, "Raw.c.nm"]))
## [1] "kldbR" "iscoR"

R.vb1i <- length(unique(rb[rb$dup.Raw & !rb$same.name & rb$ident, "Raw.c.nm"]))
as.character(unique(rb[rb$dup.Raw & !rb$same.name & rb$ident, "Raw.c.nm"]))
```

8
## 1  "flag_isco_kldbR"

(R.vb1n <- length(unique(rb[rb$dup.Raw & rb$same.name & !rb$ident, "Raw.c.nm"])))

## 1  0

(R.vb1ni <- length(unique(rb[rb$dup.Raw & rb$same.name & rb$ident, "Raw.c.nm"])))

## 1  703

(F.vb <- length(unique(rb[!rb$dup.Fin & !rb$Fin.same.name & !rb$Fin.ident, "Fin.c.nm"])))

## 1  0

(F.vbi <- length(unique(rb[!rb$dup.Fin & !rb$Fin.same.name & rb$Fin.ident, "Fin.c.nm"])))

## 1  0

(F.vbn <- length(unique(rb[!rb$dup.Fin & rb$Fin.same.name & !rb$Fin.ident, "Fin.c.nm"])))

## 1  0

(F.vbni <- length(unique(rb[!rb$dup.Fin & rb$Fin.same.name & rb$Fin.ident, "Fin.c.nm"])))

## 1  225

F.vb1 <- length(unique(rb[rb$dup.Fin & !rb$same.name & !rb$ident, "Fin.c.nm"]))
as.character(unique(rb[rb$dup.Fin & !rb$same.name & !rb$ident, "Fin.c.nm"]))

## 1  "kldb1992F"  "isco88F"

F.vb1i <- length(unique(rb[rb$dup.Fin & !rb$same.name & rb$ident, "Fin.c.nm"]))
as.character(unique(rb[rb$dup.Fin & !rb$same.name & rb$ident, "Fin.c.nm"]))

## 1  "flag_isco88_kldb1992F"

(F.vb1n <- length(unique(rb[rb$dup.Fin & rb$same.name & !rb$ident, "Fin.c.nm"])))

## 1  0

(F.vb1ni <- length(unique(rb[rb$dup.Fin & rb$same.name & rb$ident, "Fin.c.nm"])))

## 1  703
A.4. comparison summary for anchor1_DD.dta Release 2.0 and Release 3.0

<table>
<thead>
<tr>
<th>class</th>
<th>Release 2</th>
<th>Release 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ni</td>
<td>423</td>
<td>423</td>
</tr>
<tr>
<td>v</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>vn</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>vb</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbi</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbn</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbni</td>
<td>225</td>
<td>225</td>
</tr>
<tr>
<td>vb1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>vb1i</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>vb1n</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1ni</td>
<td>703</td>
<td>703</td>
</tr>
<tr>
<td>sum</td>
<td>1398</td>
<td>1402</td>
</tr>
</tbody>
</table>

Most variables are unchanged (ni, vbni, vb1ni). A small number (x, v) had been dropped or added. A number of variables (vb1, vb1i) had been (probably) renamed or recoded. A group of variables (vn) had been changed. This analysis provide no further hints for this group.

A.5. selected in-depth comparison

For illustrative purposes we present some detailed results.

```r
D1 <- cbind(Raw$kldbR, Fin$kldb1992F)
D1[D1[, 1] != D1[, 2], ]
## [,1] [,2]
## [1,] "-1" "-7"
## [2,] "-1" "-7"
```

Both variables share the same number of levels. The level '-1' has been recoded to '-7'. Changing '-1' to '-7' for two individuals takes 2 Levenshtein steps (see details in [TR-2012-003]).

```r
D2 <- cbind(Raw$dobmR, Fin$dobmF)
D2[D2[, 1] != D2[, 2], ]
## [,1] [,2]
## [1,] "5" "6"
## [2,] "8" "1"
## [3,] "6" "12"
## [4,] "10" "11"
## [5,] "11" "8"
```

Both variables share the same number of levels. $12 + 12 + 10 = 34$, $34/2 = 17$ mappings between the levels are in use. For a bijective mapping we need 5 mappings less. Coincidentally 5 individual show a changed level. It takes 7 Levenshtein steps to make the levels equal.

```r
D3 <- cbind(Raw$pageR, Fin$pageF)
D3[D3[, 1] != D3[, 2], ]
```
22 individuals have a new level for 'page'. It takes 25 Levenshtein steps to get them equal. There is no bijective mapping available.

```r
D4 <- cbind(Raw$iseiR, Fin$iseiF)
dim(D4[D4[, 1] != D4[, 2], ])
## [1] 918 2
head(D4[D4[, 1] != D4[, 2], ])
## [,1] [,2]
## [1,] "23" "5153"
## [2,] "53" "37"
## [3,] "38" "51"
## [4,] "48" "3435"
## [5,] "51" "53"
## [6,] "51" "3344"
```

918 individuals do have a different level. There is no bijective mapping.

### B. anchor2

Using the same tools we compare the data set anchor2_dd.dta from DemoDiff Release 2.0 with the same data set from DemoDiff Release 3.0.

At first we have to load the produced data collection. Release 2.0 files are denoted with R and Raw, release 3.0 files with F and Fin, respectively.
B.1. no variation

We compare the variables without variation first.

```r
R.x <- sum(!(RawNames[Rnv] %in% FinNames))
## [1] 8

F.x <- sum(!(FinNames[Fnv] %in% RawNames))
## [1] 0

# added or lost variance
(R.nA <- sum(!(RawNames[Rnv] %in% FinNames[Fnv])) - R.x)
## [1] 15

(F.nA <- sum(!(FinNames[Fnv] %in% RawNames[Rnv])) - F.x)
## [1] 37

Rnv2 <- data.frame(Rnv.tab)
Fnv2 <- data.frame(Fnv.tab)
Rnv2$nm <- gsub("R$", "", row.names(Rnv2))
Fnv2$nm <- gsub("F$", "", row.names(Fnv2))
RFnv2 <- merge(Rnv2, Fnv2, by = c("nm"), all = TRUE)
(R.ni <- sum((RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 612

F.ni <- R.ni

# changed values
(R.nB <- sum(!(RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 0

F.nB <- R.nB
(R.n <- R.nA + R.nB)
## [1] 15
```
\[(F.n <- F.nA + F.nB)\]

\[
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#
\#}
\texttt{sum(!FL.bi)}

## [1] 366

\texttt{(R.v <- sum(!((RawNames[!Rnv][!RL.bi] %in% FinNames))))}

## [1] 1

\texttt{RawNames[!Rnv][!RL.bi][!(RawNames[!Rnv][!RL.bi] %in% FinNames)]}

## [1] "isco"

\texttt{(R.vn <- sum(RawNames[!Rnv][!RL.bi] %in% FinNames))}

## [1] 369

\texttt{RawNames[!Rnv][!RL.bi][RawNames[!Rnv][!RL.bi] %in% FinNames]}

## [1] "d4"  "sex"  "ehc1p1n"  "ehc1p1g"
## [5] "ehc2p1"  "ehc3p1"  "ehc4p1"  "ehc1p2n"
## [9] "ehc1p2g"  "ehc2p2"  "ehc3p2"  "ehc4p2"
## [13] "ehc6p2"  "ehc4p1m8"  "ehc4p1m9"  "ehc4p1m10"
## [17] "ehc4p1m11"  "ehc4p1m12"  "ehc4p1m13"  "ehc4p1m14"
## [21] "ehc4p1m15"  "ehc4p1m16"  "ehc4p1m17"  "ehc2p2m8"
## [25] "ehc2p2m9"  "ehc2p2m10"  "ehc2p2m11"  "ehc2p2m12"
## [29] "ehc2p2m13"  "ehc2p2m14"  "ehc2p2m15"  "ehc2p2m16"
## [33] "ehc2p2m17"  "ehc3p2m8"  "ehc3p2m9"  "ehc3p2m10"
## [37] "ehc3p2m11"  "ehc3p2m12"  "ehc3p2m13"  "ehc3p2m14"
## [41] "ehc3p2m15"  "ehc3p2m16"  "ehc3p2m17"  "ehc27p12"
## [45] "ehc28p1"  "ehc29p1"  "ehc27p2i1"  "ehc27p2i2"
## [49] "ehc28p2"  "ehc29p2"  "ehc27p3i1"  "ehc27p3i2"
## [53] "ehc28p3"  "ehc29p3"  "ehc28p1m8"  "ehc28p1m9"
## [57] "ehc28p1m10"  "ehc28p1m11"  "ehc28p1m12"  "ehc28p1m13"
## [61] "ehc28p1m14"  "ehc28p1m15"  "ehc28p1m16"  "ehc28p1m17"
## [65] "ehc28p2m3"  "ehc28p2m4"  "ehc28p2m5"  "ehc28p2m6"
## [69] "ehc28p2m7"  "ehc28p2m8"  "ehc28p2m9"  "ehc28p2m10"
## [73] "ehc28p2m11"  "ehc28p2m12"  "ehc28p2m13"  "ehc28p2m14"
## [77] "ehc28p2m15"  "ehc28p2m16"  "ehc28p2m17"  "ehc28p3m4"
## [81] "ehc28p3m5"  "ehc28p3m6"  "ehc28p3m7"  "ehc28p3m8"
## [85] "ehc28p3m9"  "ehc28p3m10"  "ehc28p3m11"  "ehc28p3m12"
## [89] "ehc28p3m13"  "ehc28p3m14"  "ehc28p3m15"  "ehc28p3m16"
## [93] "ehc28p3m17"  "ehc16m4"  "ehc16m5"  "ehc16m6"
## [97] "ehc16m7"  "ehc16m8"  "ehc16m9"  "ehc16m10"
## [101] "ehc16m11"  "ehc16m12"  "ehc16m13"  "ehc16m14"
## [105] "ehc16m15"  "ehc16m16"  "ehc16m17"  "ehc17m5"
## [109] "ehc17m6"  "ehc17m7"  "ehc17m8"  "ehc17m9"
## [113] "ehc17m10"  "ehc17m11"  "ehc17m12"  "ehc17m13"
## [117] "ehc17m14"  "ehc17m15"  "ehc17m16"  "ehc17m17"
## [121] "ehc16"  "ehc17"  "ehc22p1n"  "ehc22p2n"
## [125] "ehc22p3n"  "ehc22p4n"  "ehc22p5n"  "ehc23p1"
## [129] "ehc23p2"  "ehc23p3"  "ehc23p4"  "ehc23p5"
## [133] "ehc24p1m"  "ehc24p2m"  "ehc24p3m"  "ehc24p4m"
## [137] "ehc24p5m"  "ehc24p6m"  "ehc24p1y"  "ehc24p2y"
## 

```
(F.v <- sum(!(FinNames[!Fnv][!FL.bi] %in% RawNames)))
## [1] 7

FinNames[!Fnv][!FL.bi][!(FinNames[!Fnv][!FL.bi] %in% RawNames)]
## [1] "flag_isco88_kldb1992" "flag_isco08_kldb2010" "ykage"
## [4] "ykid" 
## [7] "isco88"

(F.vn <- sum(FinNames[!Fnv][!FL.bi] %in% RawNames))
## [1] 359

FinNames[!Fnv][!FL.bi][FinNames[!Fnv][!FL.bi] %in% RawNames]

# [1] "d4" 
# [5] "ehc2p1" "ehc3p1" "ehc4p1" "ehc1p2n"
# [9] "ehc1p2g" "ehc2p2" "ehc3p2" "ehc4p2"
# [13] "ehc6p2" "ehc4p1m8" "ehc4p1m9" "ehc4p1m10"
# [17] "ehc4p11" "ehc4p1m12" "ehc4p1m13" "ehc4p1m14"
# [21] "ehc4p15" "ehc4p1m16" "ehc4p1m17" "ehc2p2m8"
# [25] "ehc2p2m9" "ehc2p2m10" "ehc2p2m11" "ehc2p2m12"
# [29] "ehc2p2m13" "ehc2p2m14" "ehc2p2m15" "ehc2p2m16"
# [33] "ehc3p2m17" "ehc3p2m18" "ehc3p2m19" "ehc3p2m20"
# [37] "ehc3p2m11" "ehc3p2m12" "ehc3p2m13" "ehc3p2m14"
# [41] "ehc3p2m15" "ehc3p2m16" "ehc3p2m17" "ehc3p2m18"
# [45] "ehc4p2m9" "ehc4p2m10" "ehc4p2m11" "ehc4p2m12"
# [49] "ehc4p2m13" "ehc4p2m14" "ehc4p2m15" "ehc4p2m16"
# [53] "ehc4p2m17" "ehc27p1i2" "ehc28p1" "ehc29p1"
# [57] "ehc27p2i1" "ehc27p2i2" "ehc28p2" "ehc29p2"
# [61] "ehc27p3i1" "ehc27p3i2" "ehc28p3" "ehc29p3"
# [65] "ehc27p4i1" "ehc27p4i2" "ehc28p4m" "ehc29p4m"
# [69] "ehc28p1m10" "ehc28p1m11" "ehc28p1m12" "ehc28p1m13"
# [73] "ehc28p1m14" "ehc28p1m15" "ehc28p1m16" "ehc28p1m17"
# [77] "ehc28p2m1" "ehc28p2m2" "ehc28p2m3" "ehc28p2m4"
# [81] "ehc28p2m5" "ehc28p2m6" "ehc28p2m7" "ehc28p2m8"
# [85] "ehc28p2m9" "ehc28p2m10" "ehc28p2m11" "ehc28p2m12"
# [89] "ehc28p2m13" "ehc28p2m14" "ehc28p2m15" "ehc28p2m16"
# [93] "ehc28p2m17" "ehc28p3m1" "ehc28p3m2" "ehc28p3m3"
# [97] "ehc28p3m4" "ehc28p3m5" "ehc28p3m6" "ehc28p3m7"
# [101] "ehc28p3m8" "ehc28p3m9" "ehc28p3m10" "ehc28p3m11"
# [105] "ehc28p3m12" "ehc28p3m13" "ehc28p3m14" "ehc28p3m15"
# [109] "ehc28p3m16" "ehc28p3m17" "ehc16m5" "ehc16m6"
# [113] "ehc16m7" "ehc16m8" "ehc16m9" "ehc16m10"
# [117] "ehc16m11" "ehc16m12" "ehc16m13" "ehc16m14"
# [121] "ehc16m15" "ehc16m16" "ehc16m17" "ehc16"
# [125] "ehc22p1n" "ehc22p2n" "ehc22p3n" "ehc22p4n"
# [129] "ehc22p5n" "ehc23p1" "ehc23p2" "ehc23p3"
# [133] "ehc23p4" "ehc23p5" "ehc24p1m" "ehc24p2m"
```
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ehc24p3m</td>
<td>ehc24p4m</td>
</tr>
<tr>
<td>ehc24p1y</td>
<td>ehc24p2y</td>
</tr>
<tr>
<td>ehc25p3h1</td>
<td>ehc25p4h1</td>
</tr>
<tr>
<td>ehc25p2h2</td>
<td>ehc25p3h2</td>
</tr>
<tr>
<td>ehc10k1h1</td>
<td>ehc10k2h1</td>
</tr>
<tr>
<td>ehc10k5h1</td>
<td>ehc10k6h1</td>
</tr>
<tr>
<td>hhnr</td>
<td>hh1nr</td>
</tr>
<tr>
<td>hh2a</td>
<td>hh11</td>
</tr>
<tr>
<td>hh2t0</td>
<td>hp</td>
</tr>
<tr>
<td>hpg</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hh1nr</td>
<td>hh1nr</td>
</tr>
<tr>
<td>hh2t0</td>
<td>hp</td>
</tr>
<tr>
<td>hpg</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hh1nr</td>
<td>hh1nr</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hh1nr</td>
<td>hh1nr</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hh1nr</td>
<td>hh1nr</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
<tr>
<td>hp0n</td>
<td>hp0</td>
</tr>
<tr>
<td>hpm</td>
<td>hp0g</td>
</tr>
</tbody>
</table>
Release 2.0  For 1 variables with variation but without any bijective mapping exist no variables with the same name in the Release 3.0 data. 369 variables share the name with one of the variables in Release 3.0 at least.

Release 3.0  For 7 variables with variation but without any bijective mapping exist no variables with the same name in the Release 2.0 data. 359 variables share the name with one of the variables in Release 2.0 at least.

We compare all variables pairs (with variation) which share the same name but are not connected with a bijective mapping. Those variables imply different information. DemoDiff data user should check whether they are using those variables for possible effects on their research. The number of pairs may differ from 369 and 359.

```r
results$same.name <-
  gsub("R$","",as.character(results$Raw.c.nm)) == gsub("F$","",as.character(results$Fin.c.nm))
sum(results$same.name & results$map.di>0)
## [1] 351
print(results[(results$same.name & results$map.di>0),c(2,3,5,6,7,8)],row.names=FALSE)
```

<table>
<thead>
<tr>
<th>Fin.c.nm</th>
<th>Fin.c.ls</th>
<th>Raw.c.nm</th>
<th>Raw.c.ls</th>
<th>map.di</th>
<th>ed.di.sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>d4F</td>
<td>2</td>
<td>d4R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sexF</td>
<td>2</td>
<td>sexR</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc1p1nF</td>
<td>3</td>
<td>ehc1p1nR</td>
<td>3</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>ehc1p1gF</td>
<td>3</td>
<td>ehc1p1gR</td>
<td>3</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>ehc2p1F</td>
<td>3</td>
<td>ehc2p1R</td>
<td>3</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>ehc3p1F</td>
<td>3</td>
<td>ehc3p1R</td>
<td>3</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>ehc4p1F</td>
<td>3</td>
<td>ehc4p1R</td>
<td>3</td>
<td>4</td>
<td>21</td>
</tr>
<tr>
<td>ehc1p2nF</td>
<td>2</td>
<td>ehc1p2nR</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc1p2gF</td>
<td>3</td>
<td>ehc1p2gR</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ehc2p2F</td>
<td>3</td>
<td>ehc2p2R</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ehc3p2F</td>
<td>3</td>
<td>ehc3p2R</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ehc4p2F</td>
<td>2</td>
<td>ehc4p2R</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ehc6p2F</td>
<td>2</td>
<td>ehc6p2R</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ehc4p1m8F</td>
<td>3</td>
<td>ehc4p1m8R</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m9F</td>
<td>3</td>
<td>ehc4p1m9R</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m10F</td>
<td>2</td>
<td>ehc4p1m10R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m11F</td>
<td>2</td>
<td>ehc4p1m11R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m12F</td>
<td>2</td>
<td>ehc4p1m12R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m13F</td>
<td>2</td>
<td>ehc4p1m13R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m14F</td>
<td>2</td>
<td>ehc4p1m14R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m15F</td>
<td>2</td>
<td>ehc4p1m15R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m16F</td>
<td>2</td>
<td>ehc4p1m16R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc4p1m17F</td>
<td>2</td>
<td>ehc4p1m17R</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc2p2m8F</td>
<td>3</td>
<td>ehc2p2m8R</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ehc2p2m9F</td>
<td>3</td>
<td>ehc2p2m9R</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>ehc2p2m10F</td>
<td>3</td>
<td>ehc2p2m10R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
<tr>
<td>ehc2p2m11F</td>
<td>3</td>
<td>ehc2p2m11R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
<tr>
<td>ehc2p2m12F</td>
<td>3</td>
<td>ehc2p2m12R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
<tr>
<td>ehc2p2m13F</td>
<td>3</td>
<td>ehc2p2m13R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
<tr>
<td>ehc2p2m14F</td>
<td>3</td>
<td>ehc2p2m14R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
<tr>
<td>ehc2p2m15F</td>
<td>3</td>
<td>ehc2p2m15R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
<tr>
<td>ehc2p2m16F</td>
<td>3</td>
<td>ehc2p2m16R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
<tr>
<td>ehc2p2m17F</td>
<td>3</td>
<td>ehc2p2m17R</td>
<td>2</td>
<td>1</td>
<td>1174</td>
</tr>
</tbody>
</table>
## ehc3p2m8F 2 ehc3p2m8R 3 1 2
## ehc3p2m9F 3 ehc3p2m9R 3 2 2
## ehc3p2m10F 3 ehc3p2m10R 2 1 1
## ehc3p2m11F 3 ehc3p2m11R 2 1 1
## ehc3p2m12F 3 ehc3p2m12R 2 1 1
## ehc3p2m13F 3 ehc3p2m13R 2 1 1
## ehc3p2m14F 3 ehc3p2m14R 2 1 1
## ehc3p2m15F 3 ehc3p2m15R 2 1 1
## ehc3p2m16F 3 ehc3p2m16R 2 1 1
## ehc3p2m17F 3 ehc3p2m17R 2 1 1
## ehc4p2m8F 2 ehc4p2m8R 2 2 2
## ehc4p2m9F 2 ehc4p2m9R 2 2 2
## ehc27p1i2F 9 ehc27p1i2R 8 1 2
## ehc28p1F 2 ehc28p1R 2 2 12
## ehc29p1F 3 ehc29p1R 3 4 19
## ehc27p2i1F 3 ehc27p2i1R 2 3 33
## ehc27p2i2F 16 ehc27p2i2R 15 11 34
## ehc28p2F 4 ehc28p2R 3 5 33
## ehc29p2F 4 ehc29p2R 3 1 8
## ehc27p3i1F 3 ehc27p3i1R 2 1 4
## ehc27p3i2F 6 ehc27p3i2R 5 1 4
## ehc28p3F 4 ehc28p3R 3 1 4
## ehc29p3F 3 ehc29p3R 2 1 3
## ehc28p1m8F 3 ehc28p1m8R 3 2 1
## ehc28p1m9F 3 ehc28p1m9R 3 2 2
## ehc28p1m10F 2 ehc28p1m10R 2 2 4
## ehc28p1m11F 2 ehc28p1m11R 2 2 4
## ehc28p1m12F 2 ehc28p1m12R 2 2 5
## ehc28p1m13F 2 ehc28p1m13R 2 2 9
## ehc28p1m14F 2 ehc28p1m14R 2 2 10
## ehc28p1m15F 2 ehc28p1m15R 2 2 10
## ehc28p1m16F 2 ehc28p1m16R 2 2 11
## ehc28p1m17F 2 ehc28p1m17R 2 2 12
## ehc28p2m3F 3 ehc28p2m3R 2 1 3
## ehc28p2m4F 4 ehc28p2m4R 3 3 5
## ehc28p2m5F 4 ehc28p2m5R 3 3 16
## ehc28p2m6F 4 ehc28p2m6R 3 3 16
## ehc28p2m7F 4 ehc28p2m7R 3 3 22
## ehc28p2m8F 4 ehc28p2m8R 3 5 28
## ehc28p2m9F 4 ehc28p2m9R 3 5 33
## ehc28p2m10F 4 ehc28p2m10R 3 5 33
## ehc28p2m11F 4 ehc28p2m11R 3 5 33
## ehc28p2m12F 4 ehc28p2m12R 3 5 33
## ehc28p2m13F 4 ehc28p2m13R 3 5 33
## ehc28p2m14F 4 ehc28p2m14R 3 5 33
## ehc28p2m15F 4 ehc28p2m15R 3 5 33
## ehc28p2m16F 4 ehc28p2m16R 3 5 33
## ehc28p2m17F 4 ehc28p2m17R 3 5 33
## ehc28p3m4F 3 ehc28p3m4R 2 1 2
## ehc28p3m5F 3 ehc28p3m5R 2 1 3
## ehc28p3m6F 3 ehc28p3m6R 2 1 4
## ehc28p3m7F 4 ehc28p3m7R 3 1 4
## ehc28p3m8F 4 ehc28p3m8R 3 1 4
## ehc28p3m9F 4 ehc28p3m9R 3 1 4
## ehc28p3m10F 4 ehc28p3m10R 3 1 4
## ehc28p3m11F 4 ehc28p3m11R 3 1 4
<table>
<thead>
<tr>
<th>EH</th>
<th>EH</th>
<th>R</th>
<th>Column1</th>
<th>Column2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ehc28p3m12F</td>
<td>4</td>
<td>ehc28p3m12R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc28p3m13F</td>
<td>4</td>
<td>ehc28p3m13R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc28p3m14F</td>
<td>4</td>
<td>ehc28p3m14R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc28p3m15F</td>
<td>4</td>
<td>ehc28p3m15R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc28p3m16F</td>
<td>4</td>
<td>ehc28p3m16R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc28p3m17F</td>
<td>4</td>
<td>ehc28p3m17R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc16m4F</td>
<td>3</td>
<td>ehc16m4R</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>ehc16m5F</td>
<td>3</td>
<td>ehc16m5R</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>ehc16m6F</td>
<td>4</td>
<td>ehc16m6R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc16m7F</td>
<td>4</td>
<td>ehc16m7R</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ehc16m8F</td>
<td>4</td>
<td>ehc16m8R</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ehc16m9F</td>
<td>4</td>
<td>ehc16m9R</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ehc16m10F</td>
<td>4</td>
<td>ehc16m10R</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>ehc16m11F</td>
<td>3</td>
<td>ehc16m11R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc16m12F</td>
<td>3</td>
<td>ehc16m12R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc16m13F</td>
<td>3</td>
<td>ehc16m13R</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc16m14F</td>
<td>4</td>
<td>ehc16m14R</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>ehc16m15F</td>
<td>3</td>
<td>ehc16m15R</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc16m16F</td>
<td>3</td>
<td>ehc16m16R</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc16m17F</td>
<td>3</td>
<td>ehc16m17R</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc16F</td>
<td>3</td>
<td>ehc16R</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc22p1nF</td>
<td>5</td>
<td>ehc22p1nR</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc22p2nF</td>
<td>4</td>
<td>ehc22p2nR</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>ehc22p3nF</td>
<td>4</td>
<td>ehc22p3nR</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>ehc22p4nF</td>
<td>4</td>
<td>ehc22p4nR</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>ehc22p5nF</td>
<td>3</td>
<td>ehc22p5nR</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc22p7nF</td>
<td>2</td>
<td>ehc22p7nR</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc23p1F</td>
<td>12</td>
<td>ehc23p1R</td>
<td>12</td>
<td>30</td>
</tr>
<tr>
<td>ehc23p2F</td>
<td>12</td>
<td>ehc23p2R</td>
<td>14</td>
<td>26</td>
</tr>
<tr>
<td>ehc23p3F</td>
<td>12</td>
<td>ehc23p3R</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>ehc23p4F</td>
<td>8</td>
<td>ehc23p4R</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>ehc23p5F</td>
<td>4</td>
<td>ehc23p5R</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>ehc24p1mF</td>
<td>15</td>
<td>ehc24p1mR</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>ehc24p2mF</td>
<td>15</td>
<td>ehc24p2mR</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>ehc24p3mF</td>
<td>14</td>
<td>ehc24p3mR</td>
<td>13</td>
<td>9</td>
</tr>
<tr>
<td>ehc24p4mF</td>
<td>8</td>
<td>ehc24p4mR</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>ehc24p5mF</td>
<td>5</td>
<td>ehc24p5mR</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>ehc24p6mF</td>
<td>4</td>
<td>ehc24p6mR</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc24p7mF</td>
<td>3</td>
<td>ehc24p7mR</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc24p1yF</td>
<td>45</td>
<td>ehc24p1yR</td>
<td>44</td>
<td>13</td>
</tr>
<tr>
<td>ehc24p2yF</td>
<td>42</td>
<td>ehc24p2yR</td>
<td>43</td>
<td>11</td>
</tr>
<tr>
<td>ehc24p3yF</td>
<td>24</td>
<td>ehc24p3yR</td>
<td>24</td>
<td>8</td>
</tr>
<tr>
<td>ehc24p4yF</td>
<td>11</td>
<td>ehc24p4yR</td>
<td>11</td>
<td>4</td>
</tr>
<tr>
<td>ehc24p5yF</td>
<td>6</td>
<td>ehc24p5yR</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>ehc24p6yF</td>
<td>4</td>
<td>ehc24p6yR</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>ehc24p7yF</td>
<td>3</td>
<td>ehc24p7yR</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc25p1h1F</td>
<td>5</td>
<td>ehc25p1h1R</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc25p2h1F</td>
<td>5</td>
<td>ehc25p2h1R</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>ehc25p3h1F</td>
<td>5</td>
<td>ehc25p3h1R</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>ehc25p4h1F</td>
<td>5</td>
<td>ehc25p4h1R</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>ehc25p5h1F</td>
<td>4</td>
<td>ehc25p5h1R</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>ehc25p7h1F</td>
<td>3</td>
<td>ehc25p7h1R</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>ehc25p1h2F</td>
<td>3</td>
<td>ehc25p1h2R</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>ehc25p2h2F</td>
<td>3</td>
<td>ehc25p2h2R</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>ehc25p3h2F</td>
<td>2</td>
<td>ehc25p3h2R</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>ehc25p4h2F</td>
<td>2</td>
<td>ehc25p4h2R</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>ehc26h1F</td>
<td>3</td>
<td>ehc26h1R</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>ehc26h2F</td>
<td>3</td>
<td>ehc26h2R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc10k1h1F</td>
<td>3</td>
<td>ehc10k1h1R</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>ehc10k2h1F</td>
<td>3</td>
<td>ehc10k2h1R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc10k3h1F</td>
<td>3</td>
<td>ehc10k3h1R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc10k5h1F</td>
<td>2</td>
<td>ehc10k5h1R</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>ehc10k6h1F</td>
<td>2</td>
<td>ehc10k6h1R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc10k7h1F</td>
<td>2</td>
<td>ehc10k7h1R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ehc10k1h2F</td>
<td>3</td>
<td>ehc10k1h2R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc10k2h2F</td>
<td>3</td>
<td>ehc10k2h2R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>ehc10k3h2F</td>
<td>3</td>
<td>ehc10k3h2R</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>ehc10k4h2F</td>
<td>3</td>
<td>ehc10k4h2R</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>hhnrF</td>
<td>3</td>
<td>hhnrR</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>hh1nrF</td>
<td>4</td>
<td>hh1nrR</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>hh2nrF</td>
<td>4</td>
<td>hh2nrR</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>hh1aF</td>
<td>3</td>
<td>hh1aR</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>hh2aF</td>
<td>3</td>
<td>hh2aR</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>hh1IF</td>
<td>13</td>
<td>hh1IR</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>hh2IF</td>
<td>12</td>
<td>hh2IR</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>hh1t0F</td>
<td>2</td>
<td>hh1t0R</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>hh2t0F</td>
<td>3</td>
<td>hh2t0R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>hpF</td>
<td>2</td>
<td>hpR</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>hpnrF</td>
<td>3</td>
<td>hpnrR</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>hp0F</td>
<td>2</td>
<td>hp0R</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>hps2F</td>
<td>3</td>
<td>hps2R</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>hp0nrF</td>
<td>2</td>
<td>hp0nrR</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>hp0F</td>
<td>2</td>
<td>hp0R</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>hp0gF</td>
<td>3</td>
<td>hp0gR</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>hp0ct0F</td>
<td>3</td>
<td>hp0ct0R</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>hc6h1F</td>
<td>7</td>
<td>hc6h1R</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>hc6h2F</td>
<td>8</td>
<td>hc6h2R</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>job17F</td>
<td>9</td>
<td>job17R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>job18hF</td>
<td>6</td>
<td>job18hR</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>sat2F</td>
<td>14</td>
<td>sat2R</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>sin3i1F</td>
<td>8</td>
<td>sin3i1R</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>sin3i2F</td>
<td>8</td>
<td>sin3i2R</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>sin1F</td>
<td>6</td>
<td>sin1R</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>sin2F</td>
<td>6</td>
<td>sin2R</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>sin6i1F</td>
<td>9</td>
<td>sin6i1R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>sin6i2F</td>
<td>9</td>
<td>sin6i2R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>sin6i3F</td>
<td>9</td>
<td>sin6i3R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>sin6i4F</td>
<td>9</td>
<td>sin6i4R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>sin6i5F</td>
<td>9</td>
<td>sin6i5R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>pa2mF</td>
<td>19</td>
<td>pa2mR</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>pa30F</td>
<td>3</td>
<td>pa30R</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>pa3F</td>
<td>11</td>
<td>pa3R</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>sdp1dF</td>
<td>5</td>
<td>sdp1dR</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>sdp1mF</td>
<td>16</td>
<td>sdp1mR</td>
<td>15</td>
<td>3</td>
</tr>
<tr>
<td>sdp1yF</td>
<td>29</td>
<td>sdp1yR</td>
<td>28</td>
<td>3</td>
</tr>
<tr>
<td>Group</td>
<td>Value 1</td>
<td>Value 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sdp2i1F</td>
<td>4</td>
<td>sdp2i1R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i2F</td>
<td>3</td>
<td>sdp2i2R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i3F</td>
<td>3</td>
<td>sdp2i3R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i4F</td>
<td>3</td>
<td>sdp2i4R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i5F</td>
<td>3</td>
<td>sdp2i5R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i6F</td>
<td>3</td>
<td>sdp2i6R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i7F</td>
<td>3</td>
<td>sdp2i7R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i8F</td>
<td>3</td>
<td>sdp2i8R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i9F</td>
<td>3</td>
<td>sdp2i9R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i10F</td>
<td>3</td>
<td>sdp2i10R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i11F</td>
<td>3</td>
<td>sdp2i11R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp2i12F</td>
<td>4</td>
<td>sdp2i12R</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>sdp3F</td>
<td>3</td>
<td>sdp3R</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sdp6F</td>
<td>5</td>
<td>sdp6R</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>sdp21F</td>
<td>7</td>
<td>sdp21R</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>sdp22F</td>
<td>18</td>
<td>sdp22R</td>
<td>18</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i16F</td>
<td>6</td>
<td>sdp10i16R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i17F</td>
<td>6</td>
<td>sdp10i17R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i17F</td>
<td>6</td>
<td>sdp10i17R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i18F</td>
<td>6</td>
<td>sdp10i18R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i19F</td>
<td>6</td>
<td>sdp10i19R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i19F</td>
<td>6</td>
<td>sdp10i19R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i20F</td>
<td>6</td>
<td>sdp10i20R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i21F</td>
<td>6</td>
<td>sdp10i21R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i22F</td>
<td>5</td>
<td>sdp10i22R</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i11F</td>
<td>6</td>
<td>sdp10i11R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i2F</td>
<td>6</td>
<td>sdp10i2R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i3F</td>
<td>6</td>
<td>sdp10i3R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i4F</td>
<td>6</td>
<td>sdp10i4R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i5F</td>
<td>6</td>
<td>sdp10i5R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i6F</td>
<td>6</td>
<td>sdp10i6R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i13F</td>
<td>6</td>
<td>sdp10i13R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i8F</td>
<td>6</td>
<td>sdp10i8R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i10F</td>
<td>5</td>
<td>sdp10i10R</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i11F</td>
<td>6</td>
<td>sdp10i11R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i12F</td>
<td>6</td>
<td>sdp10i12R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i15F</td>
<td>6</td>
<td>sdp10i15R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp10i14F</td>
<td>6</td>
<td>sdp10i14R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>sdp16F</td>
<td>5</td>
<td>sdp16R</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>sdp17F</td>
<td>5</td>
<td>sdp17R</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>sdp18F</td>
<td>6</td>
<td>sdp18R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>hcp1i1F</td>
<td>5</td>
<td>hcp1i1R</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>hcp1i2F</td>
<td>17</td>
<td>hcp1i2R</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>hcp2F</td>
<td>8</td>
<td>hcp2R</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>hcp3hF</td>
<td>15</td>
<td>hcp3hR</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>hcp3mF</td>
<td>18</td>
<td>hcp3mR</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>hcp4hF</td>
<td>5</td>
<td>hcp4hR</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>hcp5F</td>
<td>6</td>
<td>hcp5R</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>pa9F</td>
<td>7</td>
<td>pa9R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>pa11F</td>
<td>9</td>
<td>pa11R</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>pa13F</td>
<td>5</td>
<td>pa13R</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>pa14i1F</td>
<td>9</td>
<td>pa14i1R</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>pa14i2F</td>
<td>8</td>
<td>pa14i2R</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>pa14i3F</td>
<td>10</td>
<td>pa14i3R</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>pa14i4F</td>
<td>9</td>
<td>pa14i4R</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>pa14i5F</td>
<td>8</td>
<td>pa14i5R</td>
<td>7</td>
<td>1</td>
</tr>
</tbody>
</table>
## mageF 40  mageR 40  4  3
## k1ageF 28  k1ageR 28  2  1
## k2ageF 26  k2ageR 26  4  2
## pcobF 18  pcobR 19  3  4
## marstatF 5  marstatR 5  2  1
## npF 10  npR 10  20  23
## ncohF 7  ncohR 7  10  17
## npR 10  npR 10  20  23
## ncohR 7  ncohR 7  10  17
## meetdurF 279  meetdurR 278  13  18
## relDurF 265  relDurR 262  37  94
## cohabdurF 242  cohabdurR 249  187  134
## mardurF 191  mardurR 192  13  14
## homosexF 3  homoseXR 3  4  2
## infertileF 3  infertileR 3  4  4
## pregnantF 4  pregnantR 4  6  4
## pnkidsbioalvF 7  pnkidsbioalvR 7  12  64
## res2ndF 3  res2ndR 2  3  1175
## hhsizemrdF 13  hhsizemrdR 13  16  17
## mmarF 2  mmarR 2  2  4
## fmrdF 2  fmrdR 3  3  4
## childmrdF 9  childmrdR 8  7  8
## othmrdF 5  othmrdR 5  2  1
## hhcompF 14  hhcompR 15  21  29
## penrolF 11  penrolR 11  2  1
## pschoolF 10  pschoolR 10  6  5
## pvocatF 11  pvocatR 11  4  2
## piscedF 11  piscedR 11  6  3
## pisced2F 11  pisced2R 11  6  3
## pcasminF 13  pcasminR 13  4  2
## pyeducF 20  pyeducR 20  6  3
## egpF 129  egpR 11  264  3514
## iseiF 73  iseiR 56  451  1720
## siopsF 55  siopsR 54  431  1331
## hhincgceeF 413  hhincgceeR 414  21  118
## pcasprimF 22  pcasprimR 22  4  3
## pcasssecF 17  pcasssecR 17  2  1
## plfsF 14  plfsR 14  4  3

B.3. variation and bijective mapping

```r
# mark all identical cases
rb$ident <- (rb$ed.di.sum == 0)
# mark all not duplicated raw variables
rb$dup.Raw <- !duplicStrict(rb$Raw.c.nm)
# mark all not duplicated fin variables
rb$dup.Fin <- !duplicStrict(rb$Fin.c.nm)
# mark all equal name pairs
rb$same.name <- gsub("R$", "", as.character(rb$Raw.c.nm)) == gsub("F$", "", as.character(rb$Fin.c.nm))

rb$Raw.same.name <- (ave(rb$same.name, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.same.name <- (ave(rb$same.name, rb$Fin.c.nm, FUN = sum) > 0)
```

24
```r
rb$Raw.ident <- (ave(rb$ident, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.ident <- (ave(rb$ident, rb$Fin.c.nm, FUN = sum) > 0)
R.vb <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"]))
as.character(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"]))
## [1] "ehc24p7yR" "ehc26p7h1R"

R.vbi <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"]))
as.character(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"]))
## [1] "ehc16m1R" "ehc16m2R" "ehc16m3R" "ehc24p2m8R" "ehc24p2m9R"
## [6] "ehc10k6h2R" "ehc24p7mR" "ehc25p8h2R" "ehc25p8h3R"
## [11] "ehc10k8h2R"

R.vbn <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"]))
as.character(unique(rb[!rb$dup.Raw & rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"]))
## [1] "dpidR" "hcp4mR"

(R.vbni <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"])))
## [1] 954

R.vb1 <- length(unique(rb[rb$dup.Raw & !rb$same.name & !rb$ident, "Raw.c.nm"]))
as.character(unique(rb[rb$dup.Raw & !rb$same.name & !rb$ident, "Raw.c.nm"]))
## [1] "ehc17m4R" "ehc22p7nR" "ehc25p6h2R" "ehc25p6h3R"

R.vbi1 <- length(unique(rb[rb$dup.Raw & !rb$same.name & rb$ident, "Raw.c.nm"]))
as.character(unique(rb[rb$dup.Raw & !rb$same.name & rb$ident, "Raw.c.nm"]))
## [1] "ehc10k5h2R" "ehc10k7h2R" "klzbR"

R.vbin <- length(unique(rb[rb$dup.Raw & rb$same.name & !rb$ident, "Raw.c.nm"]))
as.character(unique(rb[rb$dup.Raw & rb$same.name & !rb$ident, "Raw.c.nm"]))
## [1] "ehc21h1R" "job18mR" "pa2yR"

(R.vbini <- length(unique(rb[rb$dup.Raw & rb$same.name & rb$ident, "Raw.c.nm"])))
## [1] 1262

(F.vb <- length(unique(rb[!rb$dup.Fin & !rb$Fin.same.name & !rb$Fin.ident, "Fin.c.nm"])))
## [1] 0

(F.vbi <- length(unique(rb[!rb$dup.Fin & !rb$Fin.same.name & rb$Fin.ident, "Fin.c.nm"]))
as.character(unique(rb[!rb$dup.Fin & !rb$Fin.same.name & rb$Fin.ident, "Fin.c.nm"]))
```

---

The code appears to be comparing and selecting unique entries based on identifiers and names in two datasets, `Raw` and `Fin`. It calculates the length of unique entries under specific conditions and categorizes them into different variables (`R.vb`, `R.vbi`, `R.vbn`, `R.vbni`, `R.vb1`, `R.vbi1`, `R.vbin`, `R.vbini`, `F.vb`, `F.vbi`). The conditions involve checking for duplicates, same names, and identifications, and the output is the count of unique entries meeting these criteria.
B.4. comparison summary for anchor2_DD.dta Release 2.0 and Release 3.0

<table>
<thead>
<tr>
<th>class</th>
<th>Release 2</th>
<th>Release 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>15</td>
<td>37</td>
</tr>
<tr>
<td>ni</td>
<td>612</td>
<td>612</td>
</tr>
<tr>
<td>v</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>vn</td>
<td>369</td>
<td>359</td>
</tr>
<tr>
<td>vb</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>vbi</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>vbn</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>vbni</td>
<td>954</td>
<td>956</td>
</tr>
<tr>
<td>vb1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>vbi1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>vb1n</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>vb1ni</td>
<td>1262</td>
<td>1260</td>
</tr>
<tr>
<td>sum</td>
<td>3246</td>
<td>3244</td>
</tr>
</tbody>
</table>

Most variables are unchanged (ni, vbni, vb1ni). A small number (x, v) had been dropped or added. Some variables (n) had lost or gained variation or changed to another constant value. A number of variables (vb,
vbi, vbn, vb1, vb1i, vb1n) had been (probably) renamed or recoded. A huge group of variables (vn) had been changed.

**B.5. selected in-depth comparison**

For illustrative purposes we present some detailed results.

```r
D1 <- cbind(Raw$pyeducR, Fin$pyeducF)
D1[D1[, 1] != D1[, 2], ]
```

```
## [,1] [,2]
## [1,] "-3.0" "-7.0"
## [2,] "14.5" "11.5"
## [3,] "-7.0" "-3.0"
```

Both variables share the same number of levels. The level '-3.0' has been recoded to '-7.0', the level '14.5' to '11.5' and the level '-7.0' to '-3.0'. This takes 3 Levenshtein steps (see details in [TR-2012-003]).

```r
D2 <- cbind(Raw$nmarR, Fin$nmarF)
D2[D2[, 1] != D2[, 2], ]
```

```
## [,1] [,2]
## [1,] "0" "1"
## [2,] "0" "1"
## [3,] "0" "1"
## [4,] "0" "1"
## [5,] "0" "1"
## [6,] "1" "0"
## [7,] "1" "0"
## [8,] "0" "1"
## [9,] "0" "1"
## [10,] "0" "1"
## [11,] "0" "1"
## [12,] "1" "2"
```

Both variables share the same number of levels. 3 + 3 + 6 = 12, 12/2 = 6 mappings between the levels are in use. For a bijective mapping we need 3 mappings less. 12 individuals show a changed level. It takes 12 Levenshtein steps to make the levels equal.

```r
D3 <- cbind(Raw$pageR, Fin$pageF)
D3[D3[, 1] != D3[, 2], ]
```

```
## [,1] [,2]
## [1,] "-7" "37"
## [2,] "26" "25"
## [3,] "-3" "25"
## [4,] "36" "38"
## [5,] "30" "29"
## [6,] "29" "30"
## [7,] "44" "46"
## [8,] "-7" "-3"
```

8 individuals have a new level for 'page'. It takes 11 Levenshtein steps to get them equal. There is no bijective mapping available.
D4 <- cbind(Raw$iseiR, Fin$iseiF)

head(D4[D4[, 1] != D4[, 2], ])

## [,1] [,2]
## [1,] "23" "22"
## [2,] "32" "33"
## [3,] "53" "31"
## [4,] "38" "56"
## [5,] "48" "50"
## [6,] "51" "48"

1128 individuals do have a different level. There is no bijective mapping.

### C. partner1

As usual we have first to load the produced data collection.

```r
cleanupStrict <- function(A) {
  return(duplicated(A) | duplicated(A, fromLast = TRUE))
}
load("..\compareR2R3\partner1\Results\compareFinRaw.RData")
```

#### C.1. no variation

We compare the variables without variation first.

```r
(R.x <- sum(!(RawNames[Rnv] %in% FinNames)))

## [1] 0

(F.x <- sum(!(FinNames[Fnv] %in% RawNames)))

## [1] 0

# added or lost variance
(R.nA <- sum(!(RawNames[Rnv] %in% FinNames[Fnv])) - R.x)

## [1] 0

(F.nA <- sum(!(FinNames[Fnv] %in% RawNames[Rnv])) - F.x)

## [1] 0
```
Rnv2 <- `data.frame`(Rnv.tab)
Fnv2 <- `data.frame`(Fnv.tab)
Rnv2$nm <- `gsub`("R\$", "", row.names(Rnv2))
Fnv2$nm <- `gsub`("F\$", "", row.names(Fnv2))
RFnv2 <- `merge`(Rnv2, Fnv2, by = c("nm"), all = TRUE)

(R.ni <- `sum`((RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 89

F.ni <- R.ni

# changed values
(R.nB <- `sum`(!(RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 0

F.nB <- R.nB

(R.n <- R.nA + R.nB)
## [1] 0

(F.n <- F.nA + F.nB)
## [1] 0

RawNames[Rnv][!(RawNames[Rnv] %in% FinNames)]
## character(0)

RFnv2[is.na(RFnv2$Fin.no.var), 1]
## character(0)

FinNames[Fnv][!(FinNames[Fnv] %in% RawNames)]
## character(0)

RFnv2[is.na(RFnv2$Raw.no.var), 1]
## character(0)

**Release 2.0** 89 variables without variation are identical in both data sets.

**Release 3.0** 89 variables without variation are identical in both data sets.
C.2. variation, but no bijective mapping

```r
c.2 <- results.bijec
RL.bi <- RawNames[!Rnv] %in% gsub("R$", "", as.character(rb$Raw.c.nm))
FL.bi <- FinNames[!Fnv] %in% gsub("F$", "", as.character(rb$Fin.c.nm))
sum(!RL.bi)
## [1] 2

sum(!FL.bi)
## [1] 2

(R.v <- sum(!((RawNames[!Rnv][!RL.bi] %in% FinNames))))
## [1] 0

(R.vn <- sum(RawNames[!Rnv][!RL.bi] %in% FinNames))
## [1] 2

RawNames[!Rnv][!RL.bi][RawNames[!Rnv][!RL.bi] %in% FinNames]
## [1] "pdobm" "pdoby"

(F.v <- sum(!((FinNames[!Fnv][!FL.bi] %in% RawNames))))
## [1] 0

(F.vn <- sum(FinNames[!Fnv][!FL.bi] %in% RawNames))
## [1] 2

FinNames[!Fnv][!FL.bi][FinNames[!Fnv][!FL.bi] %in% RawNames]
## [1] "pdobm" "pdoby"
```

**Release 2.0** 2 variables share the same name with one of the variables in Release 3.0 at least.

**Release 3.0** 2 variables share the name with one of the variables in Release 2.0 at least.

We compare all variables pairs (with variation) which share the same name but are not connected with a bijective mapping.

```r
results$same.name <- gsub("R$","",as.character(results$Raw.c.nm)) == gsub("F$","",as.character(results$Fin.c.nm))
sum(results$same.name & results$map.di>0)
```
## [1] 2

```r
print(results[(results$same.name & results$map.di>0),c(2,3,5,6,7,8)],row.names=FALSE)
```

```
# Fin.c.nm Fin.c.ls Raw.c.nm Raw.c.ls map.di ed.di.sum
# pdobmF  13  pdobmR  13   8  12
# pdobyF  38  pdobyR  38  14  36
```

C.3. variation and bijective mapping

```r
# mark all identical cases
rb$ident <- (rb$ed.di.sum == 0)
# mark all not duplicated raw variables
rb$dup.Raw <- !duplicStrict(rb$Raw.c.nm)
# mark all not duplicated fin variables
rb$dup.Fin <- !duplicStrict(rb$Fin.c.nm)
# mark all equal name pairs
rb$same.name <- gsub("R", "", as.character(rb$Raw.c.nm)) == gsub("F", "", as.character(rb$Fin.c.nm))

rb$Raw.same.name <- (ave(rb$same.name, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.same.name <- (ave(rb$same.name, rb$Fin.c.nm, FUN = sum) > 0)

rb$Raw.ident <- (ave(rb$ident, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.ident <- (ave(rb$ident, rb$Fin.c.nm, FUN = sum) > 0)

(R.vb <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"])))
```

```
# [1] 0
```

```
(R.vbi <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"])))
```

```
# [1] 0
```

```
R.vbn <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"]))
as.character(unique(rb[!rb$dup.Raw & rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"]))
```

```
# [1] "idR"
```

```
R.vbni <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"]))
as.character(unique(rb[!rb$dup.Raw & rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"]))
```

```
# [1] "pidR"
```

```
(R.vb1 <- length(unique(rb[rb$dup.Raw & !rb$same.name & !rb$ident, "Raw.c.nm"])))
```

```
# [1] 0
```

```
(R.vb1i <- length(unique(rb[rb$dup.Raw & !rb$same.name & rb$ident, "Raw.c.nm"])))
```

```
# [1] 0
```
R.vb1n <- length(unique(rb[rb$dup.Raw & rb$same.name & !rb$ident, "Raw.c.nm"]))
as.character(unique(rb[rb$dup.Raw & rb$same.name & !rb$ident, "Raw.c.nm"]))

## [1] "psd10R"

(R.vb1ni <- length(unique(rb[rb$dup.Raw & rb$same.name & rb$ident, "Raw.c.nm"])))

## [1] 137

(F.vb <- length(unique(rb[!rb$dup.Fin & !rb$Fin.same.name & !rb$Fin.ident, "Fin.c.nm"))))

## [1] 0

(F.vbi <- length(unique(rb[!rb$dup.Fin & !rb$Fin.same.name & rb$Fin.ident, "Fin.c.nm"))))

## [1] 0

F.vbn <- length(unique(rb[!rb$dup.Fin & rb$Fin.same.name & !rb$Fin.ident, "Fin.c.nm"]))
as.character(unique(rb[!rb$dup.Fin & rb$Fin.same.name & !rb$Fin.ident, "Fin.c.nm"]))

## [1] "idF"

(F.vbni <- length(unique(rb[!rb$dup.Fin & rb$Fin.same.name & rb$Fin.ident, "Fin.c.nm"))))
as.character(unique(rb[!rb$dup.Fin & rb$Fin.same.name & rb$Fin.ident, "Fin.c.nm"]))

## [1] "pidF"

(F.vb1 <- length(unique(rb[rb$dup.Fin & !rb$same.name & !rb$ident, "Fin.c.nm"))))

## [1] 0

(F.vb1i <- length(unique(rb[rb$dup.Fin & !rb$same.name & rb$ident, "Fin.c.nm"])))

## [1] 0

F.vb1n <- length(unique(rb[rb$dup.Fin & rb$same.name & !rb$ident, "Fin.c.nm"]))
as.character(unique(rb[rb$dup.Fin & rb$same.name & !rb$ident, "Fin.c.nm"]))

## [1] "psd10F"

(F.vb1ni <- length(unique(rb[rb$dup.Fin & rb$same.name & rb$ident, "Fin.c.nm"))))

## [1] 137
C.4. comparison summary for partner1_DD.dta Release 2.0 and Release 3.0

<table>
<thead>
<tr>
<th>class</th>
<th>Release 2</th>
<th>Release 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ni</td>
<td>89</td>
<td>89</td>
</tr>
<tr>
<td>v</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vn</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>vb</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbi</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbn</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>vbn1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>vb1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1i</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1n</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>vb1ni</td>
<td>137</td>
<td>137</td>
</tr>
<tr>
<td>sum</td>
<td>231</td>
<td>231</td>
</tr>
</tbody>
</table>

Most variables are unchanged (ni, vbni, vb1ni). A small number of variables (vbn, vb1n) had been (probably) renamed or recoded. A small group of variables (vn) had been changed.

D. partner2

As usual we have first to load the produced data collection.

```r
rm(list = ls())
duplicStrict <- function(A) {
  return(duplicated(A) | duplicated(A, fromLast = TRUE))
}
load("..\compareR2R3\partner2\Results\compareFinRaw.RData")
```

D.1. no variation

We compare the variables without variation first.

```r
(R.x <- sum(!(RawNames[Rnv] %in% FinNames)))
```

## [1] 0

```r
(F.x <- sum(!(FinNames[Fnv] %in% RawNames)))
```

## [1] 0

# added or lost variance

```r
(R.nA <- sum(!(RawNames[Rnv] %in% FinNames[Fnv])) - R.x)
```

## [1] 0

```r
(F.nA <- sum(!(FinNames[Fnv] %in% RawNames[Rnv])) - F.x)
```

## [1] 0
Rnv2 <- `data.frame`(Rnv.tab)
Fnv2 <- `data.frame`(Fnv.tab)
Rnv2$nm <- gsub("R$", "", row.names(Rnv2))
Fnv2$nm <- gsub("F$", "", row.names(Fnv2))
RFnv2 <- `merge`(Rnv2, Fnv2, by = c("nm"), all = TRUE)

(R.ni <- sum((RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 96

F.ni <- R.ni

# changed values
(R.nB <- sum(!(RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 0

F.nB <- R.nB

(R.n <- R.nA + R.nB)
## [1] 0

(F.n <- F.nA + F.nB)
## [1] 0

RawNames[Rnv][!(RawNames[Rnv] %in% FinNames)]
## character(0)

RFnv2[is.na(RFnv2$Fin.no.var), 1]
## character(0)

FinNames[Fnv][!(FinNames[Fnv] %in% RawNames)]
## character(0)

RFnv2[is.na(RFnv2$Raw.no.var), 1]
## character(0)

**Release 2.0** 96 variables without variation are identical in both data sets.

**Release 3.0** 96 variables without variation are identical in both data sets.
D.2. variation, but no bijective mapping

\[
\begin{align*}
\text{rb} & \leftarrow \text{results.bijec} \\
\text{RL.bi} & \leftarrow \text{RawNames[!Rnv] %in% gsub("R$", ",", \text{as.character(rb$Raw.c.nm)})} \\
\text{FL.bi} & \leftarrow \text{FinNames[!Fnv] %in% gsub("F$", ",", \text{as.character(rb$Fin.c.nm)})} \\
\text{sum}(!\text{RL.bi}) \\
\text{sum}(!\text{FL.bi}) \\
\text{(R.v} & \leftarrow \text{sum}(!\text{(RawNames[!Rnv][!RL.bi] %in% FinNames)}) \\
\text{(R.vn} & \leftarrow \text{sum}(!\text{RawNames[!Rnv][!RL.bi] %in% FinNames)}) \\
\text{RawNames[!Rnv][!RL.bi][RawNames[!Rnv][!RL.bi] %in% FinNames]} \\
\text{# [1] "pdobm" "pdoby"} \\
\text{(F.v} & \leftarrow \text{sum}(!\text{(FinNames[!Fnv][!FL.bi] %in% RawNames)}) \\
\text{(F.vn} & \leftarrow \text{sum}(!\text{FinNames[!Fnv][!FL.bi] %in% RawNames)}) \\
\text{FinNames[!Fnv][!FL.bi][FinNames[!Fnv][!FL.bi] %in% RawNames]} \\
\text{# [1] "pdobm" "pdoby"}
\end{align*}
\]

**Release 2.0**  2 variables share the name with one of the variables in Release 3.0 at least.

**Release 3.0**  2 variables share the name with one of the variables in Release 2.0 at least.

We compare all variables pairs (with variation) which share the same name but are not connected with a bijective mapping.

\[
\text{results$same.name} \leftarrow \\
\quad \text{gsub("R$","",\text{as.character(results$Raw.c.nm)}) == gsub("F$","",\text{as.character(results$Fin.c.nm)})} \\
\quad \text{sum(results$same.name & results$map.di>0)}
\]
D.3. Variation and bijective mapping

```r
# Mark all identical cases
rb$ident <- (rb$ed.di.sum == 0)
# Mark all not duplicated raw variables
rb$dup.Raw <- !duplicStrict(rb$Raw.c.nm)
# Mark all not duplicated fin variables
rb$dup.Fin <- !duplicStrict(rb$Fin.c.nm)
# Mark all equal name pairs
rb$same.name <- gsub("R$", "", as.character(rb$Raw.c.nm)) == gsub("F$", "", as.character(rb$Fin.c.nm))
rb$Raw.same.name <- (ave(rb$same.name, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.same.name <- (ave(rb$same.name, rb$Fin.c.nm, FUN = sum) > 0)
rb$Raw.ident <- (ave(rb$ident, rb$Raw.c.nm, FUN = sum) > 0)
rb$Fin.ident <- (ave(rb$ident, rb$Fin.c.nm, FUN = sum) > 0)

(R.vb <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"])))
## [1] 0

(R.vbi <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"])))
## [1] 0

(R.vbn <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"])))
## [1] 0

R.vbni <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"]))
as.character(unique(rb[!rb$dup.Raw & rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"]))
## [1] "idR" "pidR"

(R.vb1 <- length(unique(rb[!rb$ident, "Raw.c.nm"])))
## [1] 0

(R.vb1i <- length(unique(rb[!rb$ident, "Raw.c.nm"])))
```

36
## [1] 0
(R.vb1n <- length(unique(rb[dup.Raw & rb$name & !ident, "Raw.c.nm"])))
## [1] 0
(R.vb1ni <- length(unique(rb[dup.Raw & rb$name & ident, "Raw.c.nm"])))
## [1] 180
(F.vb <- length(unique(rb[!dup.Fin & !Fin.name & !Fin.ident, "Fin.c.nm"])))
## [1] 0
(F.vbi <- length(unique(rb[!dup.Fin & !Fin.name & Fin.ident, "Fin.c.nm"])))
## [1] 0
(F.vbn <- length(unique(rb[!dup.Fin & Fin.name & !Fin.ident, "Fin.c.nm"])))
## [1] 0
(F.vbni <- length(unique(rb[!dup.Fin & Fin.name & Fin.ident, "Fin.c.nm"])))
as.character(unique(rb[!dup.Fin & Fin.name & Fin.ident, "Fin.c.nm"]))
## [1] "idF" "pidF"
(F.vb1 <- length(unique(rb[dup.Fin & !name & !ident, "Fin.c.nm"])))
## [1] 0
(F.vb1i <- length(unique(rb[dup.Fin & !name & ident, "Fin.c.nm"])))
## [1] 0
(F.vb1n <- length(unique(rb[dup.Fin & name & !ident, "Fin.c.nm"])))
## [1] 0
(F.vb1ni <- length(unique(rb[dup.Fin & name & ident, "Fin.c.nm"])))
## [1] 180
D.4. comparison summary for partner2_DD.dta Release 2.0 and Release 3.0

<table>
<thead>
<tr>
<th>class</th>
<th>Release 2</th>
<th>Release 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ni</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>v</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vn</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>vb</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbi</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbn</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbni</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>vbl</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbli</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbln</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1ni</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>sum</td>
<td>280</td>
<td>280</td>
</tr>
</tbody>
</table>

Most variables are unchanged (ni, vbni, vb1ni). A small group of variables (vn) had been changed.

E. weights

As usual we have first to load the produced data collection.

```r
rm(list = ls())
duplicStrict <- function(A) {
  return(duplicated(A) | duplicated(A, fromLast = TRUE))
}
load("..\compareR2R3\weights\Results\compareFinRaw.RData")
```

E.1. no variation

We compare the variables without variation first.

```r
(R.x <- sum(!(RawNames[Rnv] %in% FinNames)))
## [1] 0

(F.x <- sum(!(FinNames[Fnv] %in% RawNames)))
## [1] 0

# added or lost variance
(R.nA <- sum(!(RawNames[Rnv] %in% FinNames[Fnv])) - R.x)
## [1] 0

(F.nA <- sum(!(FinNames[Fnv] %in% RawNames[Rnv])) - F.x)
## [1] 0
```
Rnv2 <- `data.frame`(Rnv.tab)
Fnv2 <- `data.frame`(Fnv.tab)
Rnv2$nm <- `gsub`("$", ",", `row.names`(Rnv2))
Fnv2$nm <- `gsub`("$", ",", `row.names`(Fnv2))
RFnv2 <- `merge`(Rnv2, Fnv2, by = c("nm"), all = TRUE)
(R.ni <- `sum`((RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 0
F.ni <- R.ni

# changed values
(R.nB <- `sum`(!RFnv2$Raw.no.var == RFnv2$Fin.no.var), na.rm = TRUE))
## [1] 0
F.nB <- R.nB
(R.n <- R.nA + R.nB)
## [1] 0
(F.n <- F.nA + F.nB)
## [1] 0

RawNames[Rnv][!(RawNames[Rnv] %in% FinNames)]
## character(0)
RFnv2[is.na(RFnv2$Fin.no.var), 1]
## character(0)

FinNames[Fnv][!(FinNames[Fnv] %in% RawNames)]
## character(0)
RFnv2[is.na(RFnv2$Raw.no.var), 1]
## character(0)

**Release 2.0** No change for variables without variation.

**Release 3.0** No change for variables without variation.
E.2. variation, but no bijective mapping

```r
rb <- results.bijec
RL.bi <- RawNames[!Rnv] %in% gsub("R\$", ", as.character(rb$Raw.c.nm))
FL.bi <- FinNames[!Fnv] %in% gsub("F\$", ", as.character(rb$Fin.c.nm))
sum(!RL.bi)
## [1] 0
sum(!FL.bi)
## [1] 2

(R.v <- sum(!(RawNames[!Rnv][!RL.bi] %in% FinNames)))
## [1] 0
(R.vn <- sum(RawNames[!Rnv][!RL.bi] %in% FinNames))
## [1] 0

(F.v <- sum(!(FinNames[!Fnv][!FL.bi] %in% RawNames)))
## [1] 2
FinNames[!Fnv][!FL.bi][!(FinNames[!Fnv][!FL.bi] %in% RawNames)]
## [1] "lweightDD4" "lweightPF4"
(F.vn <- sum(FinNames[!Fnv][!FL.bi] %in% RawNames))
## [1] 0
```

**Release 2.0**  No changed variables with variation but without any bijective mapping.

**Release 3.0**  For 2 variables with variation but without any bijective mapping exist no variables with the same name in the Release 2.0 data.

See the comparison table as an illustration.

```r
results$same.name <-
gsub("R\$","",as.character(results$Raw.c.nm)) == gsub("F\$","",as.character(results$Fin.c.nm))
print(results[(results$same.name),c(2,3,5,6,7,8)],row.names=FALSE)
```

<table>
<thead>
<tr>
<th></th>
<th>Fin.c.nm</th>
<th>Fin.c.ls</th>
<th>Raw.c.nm</th>
<th>Raw.c.ls</th>
<th>map.di</th>
<th>ed.di.sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>idF</td>
<td>13891</td>
<td>idR</td>
<td>13891</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>pairfamF</td>
<td>2</td>
<td>pairfamR</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>caweightF</td>
<td>535</td>
<td>caweightR</td>
<td>535</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>
E.3. variation and bijective mapping

```r
cb$ident <- (cb$ed.di.sum == 0)
cb$dup.Raw <- !duplicStrict(cb$Raw.c.nm)
cb$dup.Fin <- !duplicStrict(cb$Fin.c.nm)
cb$same.name <- gsub("R", "", as.character(cb$Raw.c.nm)) == gsub("F", "", as.character(cb$Fin.c.nm))
```

```
rb$Raw.same.name <- (ave(rb$same.name, rb$Raw.c.nm, FUN = sum) > 0)
rh$Fin.same.name <- (ave(rb$same.name, rb$Fin.c.nm, FUN = sum) > 0)
rb$Raw.ident <- (ave(rb$ident, rb$Raw.c.nm, FUN = sum) > 0)
rh$Fin.ident <- (ave(rb$ident, rb$Fin.c.nm, FUN = sum) > 0)
```

```
(R.vb <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"])))
# [1] 0
```

```
(R.vbi <- length(unique(rb[!rb$dup.Raw & !rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"])))
# [1] 0
```

```
(R.vbn <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & !rb$Raw.ident, "Raw.c.nm"])))
# [1] 0
```

```
(R.vbni <- length(unique(rb[!rb$dup.Raw & rb$Raw.same.name & rb$Raw.ident, "Raw.c.nm"])))
# [1] 0
```

```
(R.vb1 <- length(unique(rb[rb$dup.Raw & !rb$same.name & !rb$ident, "Raw.c.nm"])))
# [1] 0
```

```
(R.vb1i <- length(unique(rb[rb$dup.Raw & !rb$same.name & rb$ident, "Raw.c.nm"])))
```

```
# [1] 0
(R.vb1n <- length(unique(rb[rbdup.Raw & rbsame.name & !rbident, "Raw.c.nm"])))
# [1] 0
R.vb1ni <- length(unique(rb[rbdup.Raw & rbsame.name & rbident, "Raw.c.nm"]))
as.character(unique(rb[rbdup.Raw & rbsame.name & rbident, "Raw.c.nm"]))
# [1] "idR"  "pairfamR"  "ca1weightR"  "westR"
# [5] "ca1weightR"  "ddweightR"  "ddca1weightR"  "d1weightR"
# [9] "dica1weightR"  "lweightDD3R"  "lweightPF2R"  "lweightPF3R"

(F.vb <- length(unique(rb[!rb.dup.Fin & !rb.Fin.same.name & !rb.Fin.ident, "Fin.c.nm"])))
# [1] 0
(F.vbi <- length(unique(rb[!rb.dup.Fin & !rb.Fin.same.name & rb.Fin.ident, "Fin.c.nm"])))
# [1] 0
(F.vbn <- length(unique(rb[!rb.dup.Fin & rb.Fin.same.name & !rb.Fin.ident, "Fin.c.nm"])))
# [1] 0
(F.vbni <- length(unique(rb[!rb.dup.Fin & rb.Fin.same.name & rb.Fin.ident, "Fin.c.nm"])))
# [1] 0
(F.vb1 <- length(unique(rb[rb.dup.Fin & !rb.same.name & !rb.ident, "Fin.c.nm"])))
# [1] 0
(F.vb1i <- length(unique(rb[rb.dup.Fin & !rb.same.name & rb.ident, "Fin.c.nm"])))
# [1] 0
(F.vb1n <- length(unique(rb[rb.dup.Fin & rb.same.name & !rb.ident, "Fin.c.nm"])))
# [1] 0
(F.vb1ni <- length(unique(rb[rb.dup.Fin & rb.same.name & rb.ident, "Fin.c.nm"])))
as.character(unique(rb[rb.dup.Fin & rb.same.name & rb.ident, "Fin.c.nm"]))
# [1] "idF"  "pairfamF"  "ca1weightF"  "westF"
# [5] "ca1weightF"  "ddweightF"  "ddca1weightF"  "d1weightF"
# [9] "dica1weightF"  "lweightDD3F"  "lweightPF2F"  "lweightPF3F"
```

42
### E.4. comparison summary for weights_DD.dta Release 2.0 and Release 3.0

<table>
<thead>
<tr>
<th>class</th>
<th>Release 2</th>
<th>Release 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>n</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ni</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>v</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>vn</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbi</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbn</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vbni</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1i</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1n</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>vb1ni</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>sum</td>
<td>12</td>
<td>14</td>
</tr>
</tbody>
</table>

All old variables are unchanged (vb1ni). A small number (v) had been added.

### Literatur

[DemoDiff 2.0] Kreyenfeld, Michaela; Goldstein, Joshua; Walke, Rainer; Trappe, Heike; Huinink, Johannes (2013): Demographic Differences in Life Course Dynamics in Eastern and Western Germany (DemoDiff). GESIS Datenarchiv, Köln. ZA5684 Datenfile Version 2.0.0, http://dx.doi.org/doi:10.4232/demodiff.5684.2.0.0

[DemoDiff 3.0] Kreyenfeld, Michaela; Goldstein, Joshua; Walke, Rainer; Trappe, Heike; Huinink, Johannes (2013): Demographic Differences in Life Course Dynamics in Eastern and Western Germany (DemoDiff). GESIS Data Archive, Cologne. ZA5684 Data file Version 3.0.0, http://dx.doi.org/doi:10.4232/demodiff.5684.3.0.0
