

(参考) 今回のプロジェクトと個人情報保護制度との関係及び具体的匿名化処理の方法について

令和5年4月1日に「デジタル社会の形成を図るための関係法律の整備に関する法律」(令和三年法律第三十七号)51条に関する部分が施行され、個人情報に関する規律が個人情報保護法に一元化されました。その改正を踏まえて、今回のプロジェクトと個人情報保護制度の関係及び具体的匿名化処理の方法について以下の通り整理しました。

※ この資料に関して、不明点、疑問点等あれば、遠慮なく問い合わせてください。

## 目次

I. 個人情報保護制度に関する最近の動き .....	1
II. 今回の進め方.....	1
III. 個人情報保護法における今回提供いただく情報に関する規律.....	2
IV. 具体的匿名化処理.....	8

### I. 個人情報保護制度に関する最近の動き

- これまで、自治体における個人情報保護については、各自治体の条例において規律されていましたが、デジタル社会形成整備法(令和3年5月19日公布)により個人情報保護法が改正され、施行日である令和5年4月1日以降は、自治体における個人情報保護についても国の新たな個人情報保護法によって規律されることになりました。したがって、本プロジェクトに関しても、本法にしたがった形で情報を提供いただく必要があります。

○令和3年 改正個人情報保護法について(官民を通じた個人情報保護制度の見直し)  
(<https://www.ppc.go.jp/personalinfo/minaoshi/>)  
「デジタル社会の形成を図るための関係法律の整備に関する法律」の公布について  
令和3年2月9日に第204回通常国会に提出されました「デジタル社会の形成を図るための関係法律の整備に関する法律案」は、令和3年5月12日の国会において可決、成立し、同年5月19日に公布されました。  
デジタル社会の形成を図るための関係法律の整備に関する法律(以下「デジタル社会形成整備法」という。)により改正された個人情報保護法の施行日は、行政機関及び独立行政法人等に関する規律の規定や学術研究機関等に対する適用除外規定の見直し等(デジタル社会形成整備法第50条による改正)については公布の日から起算して一年を超えない範囲内において政令で定める日、地方公共団体に関する規律の規定(デジタル社会形成整備法第51条による改正)については公布の日から起算して二年を超えない範囲内において政令で定める日としております。

- 上記改正の施行前であった令和4年度と異なり、令和5年度以降は個人情報保護法に基づいた対応をお願いしております。各自治体が個人情報保護法上負う義務及びその適用関係については各自治体の判断となりますが、判断に当たっては、この資料を参考にしてください。

### II. 今回の進め方

- 令和5年度以降は、経年分析等の研究上の必要のため、提供いただく情報は、原則として簡易な匿名化で提供いただくことをお願いしております。特に、令和5年度以降の参加自治体については、誠

に勝手ながら、簡易な匿名化で提供いただくことが参加条件となります。

- 簡易な匿名化を経て提供いただく情報は、個人情報保護法上、各自治体にとって「個人情報」に該当します。(後に詳述)。ただし、後述する理由により、「個人情報」を、学術研究を目的とする本プロジェクトにご提供いただくことは、法的に問題がないと考えています。
- 令和 4 年度からの継続自治体については、簡易な匿名化で提供いただくことが困難であれば、施行後の匿名加工情報制度において求められる水準の匿名加工（以下、「高度な匿名化」といいます）、あるいはそれに満たない程度の加工（以下、「中間的な匿名化」といいます）を行うことを想定しています。匿名化の具体的処理方法については、IVをご覧ください。
- 高度な匿名化を経て提供いただく情報は、個人情報保護法上、各自治体にとって「個人情報」あるいは「行政機関等匿名加工情報」のいずれかに該当します。どちらとして提供いただくかは、法律上は各自治体の任意に委ねられていますが、後者として提供いただく場合のプロセスが双方にとって煩雑になるため、本プロジェクトでは、「個人情報」としての提供をお願いしています（Ⅲで詳述）。

### Ⅲ. 個人情報保護法における今回提供いただく情報に関する規律

#### 1. 各自治体に対する規律

(1) 今回提供いただく情報が各自治体にとって何に当たるか

- 現行法において、「仮名加工情報」(2条5項)あるいは「匿名加工情報」(2条6項)の作成に関する規律を含む第四章は、各自治体に対しては適用除外となっています(16条2項、5項、6項)。
- したがって、行政機関が、個人情報保護法上の「個人情報」に対して匿名化を行い、かつ、その匿名化の水準が「仮名加工情報」(2条5項)あるいは「匿名加工情報」(2条6項)に求められる水準に達しているとしても、それらが個人情報保護法上の「仮名加工情報」あるいは「匿名加工情報」としての規律を受けることはありません。
- それらは「個人情報」のままであるか、あるいは、法第五節が定める諸要件を満たす場合には、「行政機関等匿名加工情報」として取り扱うことができます。

## (2) 行政機関等匿名加工情報に関する規律

- 個人情報に対する加工が、「特定の個人を識別することができないように及びその作成に用いる保有個人情報を復元することができないようにするために必要なものとして個人情報保護委員会規則で定める基準に従い」(106条1項)行われた場合には、それを「行政機関等匿名加工情報」として取り扱うことができます。
- ただし、上記のような加工を行ったとしても、必ず当該情報を行政機関等匿名加工情報として扱わなければならないわけではなく、法的には「個人情報」として取り扱うことも可能です。
- 「行政機関等匿名加工情報」は、「個人情報」と異なり(のちに詳述)、利用目的外の学術目的での第三者提供が認められていません(109条2項)。そのため、提供いただく情報を「行政機関等匿名加工情報」として扱う場合には、概ね以下のような手続きを経る必要があります。
  - ① 税務データを加工した行政機関等匿名加工情報の個人情報ファイル簿への記載(117条)
  - ② 当該行政機関等匿名加工情報の利用に対する提案(118条)
  - ③ 地方公共団体による提案の審査(114条)
  - ④ 契約の締結(115条)+手数料の支払(119条3項)
- この手続きは双方にとって煩雑なため、本プロジェクトでは、提供いただく情報は、法的には「個人情報」として取り扱っていただくことをお願いしております。

## (3) 個人情報に関する規律

- 個人情報保護法69条は、個人情報の利用(匿名化などの加工を含む)及び提供を制限しています。ただし、「学術研究の目的のために保有個人情報を提供するとき」には、保有個人情報の利用及び提供を例外的に許容しています(同条2項4号)。

### (利用及び提供の制限)

第六十九条 行政機関の長等は、法令に基づく場合を除き、利用目的以外の目的のために保有個人情報を自ら利用し、又は提供してはならない。

2 前項の規定にかかわらず、行政機関の長等は、次の各号のいずれかに該当すると認めるときは、利用目的以外の目的のために保有個人情報を自ら利用し、又は提供することができる。ただし、保有個人情報を利用目的以外の目的のために自ら利用し、又は提供することによって、本人又は第三者の権利利益を不当に侵害するおそれがあると認められるときは、この限りでない。

一～三 (省略)

四 前三号に掲げる場合のほか、専ら統計の作成又は学術研究の目的のために保有個人情報を提供するとき、本人以外の者に提供することが明らかに本人の利益になるとき、その他保有個人情報を提供することについて特別の理由があるとき。

- 本プロジェクトの目的は、政策現場における EBPM の推進とアカデミアにおける実証研究の発展であり、科学研究費助成事業から助成を受けている（研究種目：学術変革領域研究(B) 研究領域：自治体データ整備）ため、本プロジェクトが正当な学術研究であることは明らかだと考えています。したがって、本プロジェクトに対して提供するために必要な限度において、「個人情報」を匿名化したり提供したりすることは、法 69 条 2 項 4 号によって許容され则认为します。
- 同号に基づく提供を行った際に、各自治体は、「必要があると認めるときは、保有個人情報の提供を受ける者に対し、提供に係る個人情報について、その利用の目的若しくは方法の制限その他必要な制限を付し、又はその漏えいの防止その他の個人情報の適切な管理のために必要な措置を講ずることを求めるものとする。」（70 条）とされています。  
※ 「必要な制限」又は「必要な措置」としては、具体的には、以下のような内容が想定されています。

個人情報の保護に関する法律についてのガイドライン（行政機関等編）32 頁

「必要な制限」又は「必要な措置」としては、利用の目的又は方法の制限のほか、提供に係る保有個人情報の取扱者の範囲の限定、第三者への再提供の制限又は禁止、消去や返却等利用後の取扱いの指定、提供に係る保有個人情報の取扱状況に関する所要の報告の要求等が考えられる。

- 東京大学政策評価研究教育センターでは、70 条及びガイドラインの明示する制限及び要求に対応する形での取り扱いをすることを内規で定めております（2（3）で詳述）。これは、個人情報保護法上、個人情報取扱事業者が負う義務に準じたものです。
- したがって、本プロジェクトに関連して本学に対して「個人情報」を提供いただいても法的な問題はないと認めます。

## 2. 東京大学評価研究教育センターに対する規律

（1）今回提供いただく情報が東京大学評価研究教育センターにとって何に当たるか

- 今回提供いただく情報は、少なくとも簡易な匿名化を経ているため、本学にとっては「個人情報」に当たりません。
- また、行政機関が作成する情報は、「仮名加工情報」あるいは「匿名加工情報」としての規律を受けません。
- したがって、本学にとって、提供いただく情報は、「個人関連情報」に該当します（2 条 7 項）。

## (2) 個人関連情報に関する個人情報保護法上の規律

- 情報を提供いただいた場合、本学は「個人関連情報取扱事業者」に該当します（16条7項）。
- 個人関連情報取扱事業者に対する規律は、「第三者が個人関連情報……を個人データとして取得することが想定されるとき」（31条1項）の制限のみです。
- 本プロジェクトでは、「第三者が個人関連情報……を個人データとして取得することが想定されるとき」に該当するかどうかを問わず、提供いただいたデータは、原則として、第三者に提供しません。
- ただし、研究成果を発表するための学術雑誌から、データから研究結果が導けることを確認する目的（再現性目的）で、データの提供を求められることがあります。近時、様々な学術的問題を背景として、再現性目的でデータを提供しない限りは論文の投稿を受け付けない、あるいは掲載まで至らない雑誌が急増しています。そのため、本研究の結果を論文として適切な媒体で公表するためには、再現性目的に限定して、学術雑誌に対してデータを提供できることが必要不可欠です。

そこで、この場合に限って、例外として、センター長がデータ提供先や提供先における守秘義務・情報管理体制を確認した上で、再現性目的に必要な限りで、提供いただいたデータを学術雑誌側に提供します。当然ながら、学術雑誌側はこれ以外の目的でデータを用いることはなく、また、当該雑誌への掲載プロセスに無関係な研究者がそれにアクセスすることはありません。
- この場合には、提供したデータを提供元である自治体に対しては、センターから、提供を行った旨及び提供先に関する情報を遅滞なく通知します。
- なお、学術雑誌側が、センターが提供するデータを「個人データとして取得すること」は想定されませんので、この場合でも、31条に基づく提供には該当しません。

## (3) 提供いただいたデータに対する東京大学評価研究教育センターによる規律

- 東京大学評価研究教育センターでは、「自治体税務データ活用プロジェクトにおけるデータ取扱規則」及び「自治体税務データ活用プロジェクトにおける安全管理措置等に関する規程」において、提供いただいたデータに対して個人情報取扱事業者準じた取り扱いを行う義務を定めています。
- 例えば、ガイドラインで挙げられた「必要な措置」及び「必要な制限」とデータ取扱規則は、以下のように対応しています。

※ ここで挙げているのは規則・規程の一部ですので、別途、全体をご覧ください。

## ① 利用の目的の制限

### (データの利用目的)

第2条 政策評価研究教育センター（以下「センター」という。）は、本プロジェクトにおいて自治体から提供された個人又は法人に関するデータ（以下「自治体提供データ」という。）を、センターが行う税や経済等に関する学術研究としてセンター長が定める利用目的の達成に必要な範囲を超えて取り扱ってはならない。

## ② 利用方法の制限

### (識別行為の禁止)

第3条 センターは、自治体提供データのうち、個人に関するデータを取り扱うに当たっては、当該自治体提供データの作成に用いられた個人情報に係る本人を識別するために、当該個人情報から削除された記述等若しくは個人識別符号（個人情報の保護に関する法律（平成15年法律第57号）第2条第2項に規定する個人識別符号をいう。）に関する情報を取得し、又は当該自治体提供データを他の情報と照合してはならない。

2 センターは、自治体提供データのうち、法人に関するデータで、データ提供時までに自治体が当該法人を識別してはならない旨の意思を表示したものを取り扱うに当たっては、当該自治体提供データの作成に用いられた法人情報に係る当該法人を識別するために、当該法人情報から削除された記述等に関する情報を取得し、又は当該自治体提供データを他の情報と照合してはならない。

### (データの保存と廃棄)

第5条 センター及びプロジェクト従事者は、自治体提供データ（以下「原データ」という。）及びそれを複製したもの（以下「複製データ」という。）に変更が加えられ、又は漏えい、滅失若しくはき損しないよう管理しなければならない。

- ③ 提供に係る保有個人情報の取扱者の範囲の限定
- ④ 第三者への再提供の制限又は禁止

(データの取扱者)

第4条 自治体提供データについては、本プロジェクトに従事する者としてセンター長が定める者(以下「プロジェクト従事者」という。)のみが取り扱うものとし、その他の者に提供してはならない。

2 前項の規定にかかわらず、プロジェクト従事者は、学術雑誌等における再現性目的での提供が求められる場合には、センター長の許可を得て、自治体提供データを第三者に提供できるものとする。

3 前項の場合には、センターは、遅滞なく、プロジェクト従事者が第三者に提供した自治体提供データの提供元である自治体に対して、提供を行った旨及び提供先に関する情報を通知しなければならない。

- ⑤ 消去や返却等利用後の取扱いの指定

(データの保存と廃棄)

第5条 ……

3 プロジェクト従事者は、複製データが本プロジェクトの研究に不要になった場合は、速やかに当該複製データを廃棄しなければならない。

- ⑥ 提供に係る保有個人情報の取扱状況に関する所要の報告

(利用状況の報告)

第6条 センターは、自治体提供データを提供した自治体(以下「参加自治体」という)に対し、参加自治体により提供されたデータが本取扱規則の条件に適合して適切に利用されていることを証するために、データを利用する期間において毎年度1回、利用状況の報告をするものとする。

2 参加自治体は、前項に基づく報告が自治体提供データの利用状況を検証するのに十分ではないと判断した場合、センターの承諾を得て、センターの事業所において、センターによる自治体提供データの利用状況の監査を実施することができるものとする。

#### IV. 具体的匿名化処理

本プロジェクトでは、簡易な匿名化で提供いただくのが原則です。特に、令和5年度以降の参加自治体については、誠に勝手ながら、簡易な匿名化で提供いただけることが参加条件となります。

ただし、令和4年度からの継続自治体については、簡易な匿名化で提供いただくことが困難であれば、高度な匿名化あるいは中間的な匿名化を選択していただくことも可能です。

簡易な匿名化では、匿名加工情報制度において求められる水準の匿名加工とは必ずしも言えないものの、氏名、個人番号（いわゆるマイナンバー）、住所等の削除や、宛名番号、世帯番号のハッシュ化（後述）等を行います。高度な匿名化はより匿名性が高いものとなりますが、データの有用性は低下します。

「簡易な匿名化」は黒丸数字①～③の流れ、「高度な匿名化」は丸数字①～⑦の流れで匿名化処理を行うことを想定しています。「中間的な匿名化」では黒丸数字①～③に任意の丸数字①～⑦を追加した上で匿名化処理を施します。

##### （事前処理）

①① 氏名、個人番号（いわゆるマイナンバー）、住所等の事前削除

##### （匿名化処理）

① 特異な世帯の世帯番号の秘匿

①② 宛名番号、世帯番号のハッシュ化

②③ 宛名番号をキーにして、複数年のデータの結合

④ 所得、賦課額等のトップコーディング

③⑤ 生年月日の月単位への丸め

⑥ 生年月、性別等について、「3-匿名性」を判定し、「3-匿名性」を満たさないものは、「3-匿名性」を満たすまで秘匿

⑦ 各世帯番号（のハッシュ値）について、ランダムサンプリング

## 【各匿名化処理の解説】

### ① 特異な世帯の世帯番号の秘匿

特異な世帯は、それだけで個人識別性が高くなります。  
そこで、特異な世帯<sup>1</sup>の世帯番号<sup>2</sup>を秘匿します。

### ①② 宛名番号、世帯番号のハッシュ化

宛名番号、世帯番号については、ハッシュ化<sup>3</sup>を行います。

これらは通常、一般に知りうるものではなく、知りえたとしてもそこからすぐに個人情報にアクセスできるものでもありませんが、自治体内で様々なシステムを通じて横断的に使われていることもあり、万が一、流出したときの影響が大きいため、ハッシュ化を行います。年度をまたがる個人の紐付けは、このハッシュ化した宛名番号により行います。

ハッシュ化の方法としては、鍵付きハッシュ関数を用いるものとし（鍵となる秘密の文字列は自治体限りとしていただきます）、事務局レポート p. 21 において利用が推奨されている CRYPTREC（Cryptography Research and Evaluation Committees）により公開されている電子政府推奨暗号リスト<sup>4</sup>において挙げられているハッシュ関数のアルゴリズムを用います。

### ②③ 宛名番号をキーにして、複数年のデータの結合

宛名番号<sup>5</sup>をキーにして、複数年のデータを個人単位で結合します。

「II 高度な匿名化」の場合は、生年月、性別、郵便番号（郵便番号がデータセットに入っている場合。⑥において同じ。）が年度によって異なる場合は、最も古いものを優先します。自治体内で転居した場合は、年度によって郵便番号が異なることがありますが、ある郵便番号の場所から別の郵便番号の場所に転居したという情報は、個人識別性が高いため、最も古い郵便番号を保持する（転居情報を隠す）ことによって、匿名性を確保します。

<sup>1</sup> 「特異な世帯」の基準は公表していません。プロジェクト参加確定後、お示しします。

<sup>2</sup> 自治体内において世帯を一意に識別するために付番した番号のことです。

<sup>3</sup> ハッシュ化とは、元のデータから一定の計算手順に従ってハッシュ値と呼ばれる規則性のない値を求め、その値によって元のデータを置き換えることにより、データを不可逆的に別の形に変える方法をいいます。

例としては以下のようなイメージです。

宛名番号	ハッシュ化した宛名番号
12345678	aks;ldfjpawnefdoiewadlksfdajf
23456789	k3298refcsna3489elsjodslerire
34567890	kjlz-9erj8ufdjiawe8re;ofjwloe

ここで、「aks;ldfjpawnefdoiewadlksfdajf」から、元の「12345678」を復元することは不可能です。そのため、ハッシュ化を行えば当方が宛名番号を知りえないまま、年度をまたぐ個人の紐付けが可能となります。

<sup>4</sup> CRYPTREC 暗号リスト（電子政府推奨暗号リスト）（<http://www.cryptrec.go.jp/list.html>）。

<sup>5</sup> 自治体内において個人、法人を一意に識別するために付番した番号のことです。自治体によって、「個人コード」など他の名称で呼ばれることもあります。「個人番号」、「住記個人番号」と呼ばれることもありますが、番号法に基づく「個人番号」（いわゆるマイナンバー）とは異なります。

#### ④ 所得、賦課額等のトップコーディング

所得、賦課額、控除額等の金額については、同じ性別・年代ごとに 0.5%を対象として（ただし、この数が 10 人に満たない場合は 10 人を対象とします。以下同じ）、トップコーディングを行います。トップコーディングとは、加工対象に含まれる情報のうち、特に数値の大きいグループについてまとめる処理のことです。

今回は、同じ性別・年代（西暦の生年の上 3 桁）ごとに各項目（例：所得であれば所得、賦課額であれば賦課額）の上位 0.5%を、同じ値（上位 0.5%の平均値）に変換します。これにより、例えば、「男性・1970 年代生まれ」が 1 万人いる自治体であれば、上位 50 人が同じ所得となり、人数の少ない高所得者層であっても高所得であるからという理由から個人を特定するのが困難になります。

この結果、データセット中に男女それぞれで 1930 年代～2020 年代生まれが少なくとも 10 人ずついるとすると、 $10 \text{ 人} \times 2 \text{ (男女)} \times 10 \text{ (年代)} = 200 \text{ 人}$ が少なくともトップコーディングされることとなり、人口 4 万以上の場合はこれより多くの人（少なくとも人口の 0.5%）がトップコーディングの対象となります。

#### ③⑤ 生年月日の月単位への丸め

生年月日については、「日」の情報を落として、月単位に丸めます。このときの丸め方としては、生年月日の前日の年と月に丸めます。

例えば、2001 年 1 月 1 日生まれの人と 2001 年 1 月 2 日生まれの人がいれば、前者は 2021 年 1 月 1 日時点で 20 歳であり、後者は 19 歳です。行政の取扱い上、1 日時点の年齢で判断されることが多いため（学齢等）、前者については前日（2020 年 12 月 31 日）の年と月である「2020 年 12 月」に、後者については前日（2021 年 1 月 1 日）の年と月である

「2021 年 1 月」に丸めます。

もっとも、生まれ月まで入っていると、個人を特定するリスクが高くなりますが、この問題については⑥の処理で対応します。

#### ⑥ 生年月、性別、郵便番号について、「3-匿名性」を判定し、「3-匿名性」を満たさないものは、「3-匿名性」を満たすまで秘匿

データセットにある変数のうち、生年月、性別、郵便番号については、他者が容易に情報を得ることができることが多く、これらから個人を特定するリスクを避ける必要があります。「3-匿名性」<sup>6</sup>が確保されれば、生年月、性別、郵便番号が同じ個人が少なくとも 3 人（自分を除けば 2 人）いることが保証されるので、生年月、性別、郵便番号の情報を得ても、当該個人である確率は 1/3 以下となります（さらに、

<sup>6</sup> 「k-匿名性を満たす」とは、対象となるデータセット内に、同じ準識別子（今回の場合は、生年月、性別、郵便番号）の組合せを持つデータが k 件以上存在することをいいます。事務局レポート p.32 では、「匿名加工情報は、……一般公開されるものではないから、上記で準識別子とされている情報の項目について、匿名加工情報データベース等との関係で  $k \geq 2$  となるように加工することは必ずしも求められない。ただし、匿名加工情報が第三者に提供される態様や利用形態を考慮した上で、必要に応じてこのような考え方を取り入れることが望ましい。」としています。

後述の⑦の処理を行うことにより 1/6 以下となります)。

しかし、例えば、特に高齢である方、過疎地域に住む方などは、「3-匿名性」を満たさないことがあります。その場合、「3-匿名性」を満たすよう、生年月、性別、郵便番号のいずれかを秘匿<sup>7</sup>します。

秘匿の順序は、以下のとおりとします。

1. 生年月を四半期 (1~3 月、4~6 月、7~9 月、10~12 月) にする
2. 郵便番号の 7 桁目を秘匿
3. 郵便番号の 6 桁目を秘匿
4. 郵便番号の 5 桁目を秘匿
5. 郵便番号の 4 桁目を秘匿
6. 郵便番号を秘匿
7. 生年月を半期 (1~6 月、7~12 月) にする
8. 生年月の月を秘匿
9. 生年を 5 年単位で丸め
10. 生年の下一桁を秘匿
11. 生年を秘匿

これにより、すべての個人が「3-匿名性」を満たすこととなります。

#### ⑦ 各世帯番号 (のハッシュ値) について、ランダムサンプリング

⑥で得られたデータセットのうち各世帯番号 (のハッシュ値) について、50%の確率でサンプルに入れるかどうかをランダムで決めます (ランダムサンプリング)<sup>8</sup>。

これは、個人の特特定ができない状況でも、個人の情報が得られてしまうリスクに対処するためです。

例えば、「1932 年 1 月生まれ・女性・郵便番号 9876543」が 3 人いる場合、これらの情報からは 3 人のうち誰かを特定できず、「3-匿名性」を満たします。しかし、データセットの中で、「1932 年 1 月生まれ・女性・郵便番号 9876543」の 3 人の所得が全員「0 円」だったとすると、3 人の誰かは特定できなくても、所得が「0 円」であることは分かっています。

そこで、50%のランダムサンプリングを行うことにより、個人がデータセットに入っている確率が、個人がデータセットに入っていない確率を上回らないようにします。そうすることにより、上記のような推論が成り立つ可能性が、上記のような推論が成り立たない確率を上回らないこととなります。

また、個人単位ではなく、世帯単位でサンプリングを行うのは、1 世帯内でサンプルに入っている構成員と入っていない構成員がないようにするためです。

なお、50%は、あくまで確率であるため、サンプル数が元のデータセットのレコード数のちょうど 50% に一致しない可能性があります。例えば、人口 10 万ちょうどの自治体について⑦の処理を行った場合、サンプル数は平均して 5 万ちょうどのようになりますが、場合によっては 49892 人になったり 50006 人になったりします。

<sup>7</sup> 空欄に置き換えるか、別の値に置き換えますが、いずれにせよ、「3-匿名性」が満たされるまでこの処理を行います。

<sup>8</sup> この際、ある個人が年度間で異なる世帯番号を持つことが考えられるため、データセット中のいずれの年度でも、1 世帯内でサンプルに入っている構成員と入っていない構成員がないように処理します。