

平成 20 年 5 月 27 日

「統計データの二次利用促進に関する研究会」報告書

(中間取りまとめ骨子)

はじめに

報告書を取りまとめるに至った経緯。

基本的な考え方

統計法の施行に伴い、統計データの二次利用のための業務は各府省の行政活動の一環として適切に実施されることが必要。研究会では各府省が整合性を確保しながら適切に当該業務を実施していくための運用面、技術面での諸課題を検討、整理。

二次利用の運用手続面について

1 利用目的の範囲

統計法第 34 条及び第 36 条では、「学術研究の発展に資すると認める場合その他の総務省令で定める場合」にのみ、「委託による統計の作成等」(以下「オーダーメイド集計」という。)や「匿名データ」の提供ができると規定。

ア. 「公益性」を求める理由

(1) オーダーメイド集計の場合

国民の共有財産である行政資源を費やすこと等による。

(2) 匿名データの提供の場合

秘匿化されても個別情報の提供であり、目的を問わずに提供すると調査対象者の信頼を損ね、統計の真実性確保に支障のおそれがあることによる。

イ. 「学術研究の発展に資すると認める場合」

(1) 大学・研究機関における場合

教授、准教授、講師、助教、ポストドクター、大学院生等が学術を目的として研究活動を行う場合を想定。しかし、公益性確保のためには、研究成果が公表され、社会に還元されることを要件とすべき。

(2) 大学・研究機関以外における場合

営利企業に属する者の企業活動の一環としての研究も、学術的な研究で、その成果が社会に還元される場合、要件に該当。しかし、研究成果が企業の業務資料や特定顧客へのレポートである場合、要件に非該当。学術研究に営利目的が一部含まれる場合でも、研究成果が社会に還元される場合、

要件に該当。

ウ．「その他総務省令で定める場合」

大学等での教育目的等を規定すべき。この場合には、指導教授等の指導監督の下で、かつ使用場所は教室内や研究室に限るなどの条件を付すことが必要。

教育目的等で使用する場合の条件（申出書、審査）等は、今後の研究会の議論を踏まえ追加する予定。

2 利用の申出

ア．利用申出書

統計法第 34 条及び第 36 条では、申出事項は総務省令で定めることを規定。申出書様式の府省間での統一が必要。申出事項を以下に提示。

(1) オーダーメイド集計の場合

- 申出者（所属、氏名、連絡先）
- 使用する統計調査の名称及び年次
- 利用目的（研究計画名、具体的な研究内容、研究成果の取扱い（○○学会で発表、○○誌に掲載等））
- 集計等の作業委託内容（集計表様式を添付）
- 共同研究者の氏名及び所属
- 作成した集計表等の提供を受ける方法

(2) 匿名データの提供の場合

- 申出者（所属、氏名、連絡先）
- 使用する統計調査の名称及び年次
- 利用目的（研究計画名、具体的な研究内容、研究成果の取扱い（○○学会で発表、○○誌に掲載等））
- 使用場所
- データの管理方法（データの保存場所・保存方法、パソコンの使用環境（インターネット非接続等））
- データを取扱う者の範囲（共同研究者、外部委託する場合の委託先等）
- 匿名データの提供を受ける方法

イ．利用申出者

自分自身が責任を持って研究を行い得る者。大学等の利用申出者を以下のよう

に例示。
教授等の指導の下で、大学院生や学部生が研究に携わる場合、教授等共同研究として、教授等と大学院生が研究を行う場合、両者

個人として大学院生等が研究を行う場合(学部生の身分を理由とした排除

は困難)、大学院生等

なお、匿名データの提供の場合、利用申出者の本人確認を十分に行うことが必要。

ウ．利用申出の受付窓口

利用申出書は、調査票情報を保有する府省に対して提出。

エ．利用申出の時期

受付から提供までの期間を明確化。年間の受付時期・期間を年数回に限定。

3 利用目的の審査

研究成果の公表方法の明確化と「公益性」の有無がポイント。

(1) 「公益性(学術研究であること)を示す書類」

次の書類を想定。

- 研究計画書
- 所属がある場合は所属を証明する書類(指導教授や大学、学会からの推薦状、大学に籍があることを証明する書類等)
- 公表の方法に係る書類(発表予定の学会・大会の名称や活動内容、提出する研究計画書への掲載予定の学術誌、機関紙、専門誌、業界誌等の記述等)
- その他(研究費を確保するために作成した既存の書類、利用者の著書・論文の一覧等)

(2) 審査内容

依頼者の属する機関が外形的にみて研究を行っている機関か
研究の成果が社会に還元されるか
提供したデータの管理体制や管理方法に信頼が置けるか
研究内容と利用される統計調査の関係が適切であるか

(3) 審査の実施主体

以下の理由から審査は当該統計調査の実施者が行うことが適当。

-) 利用目的の審査は研究内容の重要度を評価するものではない
-) 調査対象者の信頼確保のため、調査実施者は研究内容を確認すべき
-) 統計調査の内容や匿名データの詳細を調査実施者が把握

公平性や透明性を確保する観点から、申出に応じなかった理由等を申出者に通知するとともに、統計委員会に報告。

4 審査結果に対して不服があった場合の対応方法

行政手続法第2条第2号や行政不服審査法第2条第1項の規定する「処分」

に該当せず、法的な救済制度はないが、二次利用に関する政府統計全体の相談・苦情等の窓口機能を設けるべき。

5 集計結果表等の著作権に関する留意事項

申出者や集計事務の受託業者の創意工夫で結果表が作成されると著作権が発生。成果物の著作権は調査実施者側に帰属する旨を契約で明確化すべき。また、申出者は著作権を主張しないことを申出受理の条件に入れるべき。

6 その他運用面での留意事項

- (1) 提供窓口は、匿名データ等を作成する府省毎に置く。
- (2) 提供手段は、電子媒体の郵送、直接受渡。集計表はオンラインも可。
- (3) 匿名データの利用終了後、破壊又は返却を確認できる仕組みを構築。また、オーダーメイド集計で提供した集計表等や匿名データから作成した集計表等の利用は申出書に記載した利用目的の範囲に限定すべき。
- (4) 匿名データを作成する統計調査は、各府省がニーズを勘案して判断。いつごろどのような匿名データが作成されるか分かる情報を公表すべき。また、政府統一的なニーズ把握の仕組みを構築すべき。
- (5) 匿名データは、枠組みとしては外国の研究者に対しても提供可能。
- (6) レプリカデータは定義の明確化の上で作成・提供について検討。
- (7) オンサイト利用は目的外利用のガイドライン検討において制度化し、運用方法を定めるべき。

二次利用の技術的な課題について

1 オーダーメイド集計における集計結果表の秘匿処理

集計結果表の秘匿処理の方法として一律の基準を設定することは困難。ガイドラインでは、現在各府省が統計調査結果を公表する際に実施している秘匿処理の事例（別紙1）を参考に、目安となるものを提示。

2 匿名データの作成

(1) 匿名データの作成対象

原則として、世帯に限らず事業所関係の統計調査も含め対象とすべき。但し、匿名化できない調査は、オーダーメイド集計やオンサイト利用で対応すべき。

(2) 匿名処理の方法

匿名化の方法は、調査ごとに決定。ガイドラインには秘匿処理の技法（別紙2）や匿名化の目安（別紙3）を盛り込む。

3 秘匿処理の審査

統計委員会の審査を効率的、効果的に実施するためチェックリストが必要。各府省間で統一化された様式で、各府省が所管の統計調査ごとに作成。チェックリストのイメージは別紙4及び5。

4 秘匿処理の実施機関

秘匿化という専門的・技術的処理を効率的に実施するため、ある程度秘匿処理の実施機関を特定化し、業務委託を検討すべき。

別紙1：現在、各省が実施している集計結果表の秘匿処理の事例

別紙2：秘匿処理について

別紙3：匿名化の基準（目安）

別紙4：匿名データのチェックリスト（案） 全国消費実態調査を例として

別紙5：匿名データのチェックリスト（案） 就業構造基本調査を例として

現在、各省が実施している集計結果表の秘匿処理の事例

各府省において実施している集計結果表の秘匿処理については、概ね下表のとおり。

調査対象	標本	全数
事業所等	<p>客体数が少ない場合、結果を非表示（“ x ”等に置換え）（賃金引上げ等の実態に関する調査等）</p> <p>その他、合計値からの引き算により秘匿対象が判明する場合は、二次秘匿処理（サービス業基本調査）</p> <p>客体数が3未満の場合、客体数は表章するが経営に係る項目は非表示（農林水産関係の統計全般）</p> <p>事業所数が一定数以下でかつ従業者数が一定数以下の場合非表示（屋外労働者職種別賃金調査等）</p> <p>労働者数を10人単位で表章（賃金構造基本調査）</p>	<p>客体数が少ない場合、結果を非表示（“ x ”等に置換え）</p> <p>合計値からの引き算により秘匿対象が判明する場合は、二次秘匿処理（工業統計調査、商業統計調査、学校教員統計調査等）</p>
世帯	<p>表章単位の丸め（1000世帯、万人等）（労働力調査、国民生活基礎調査等）</p>	<p>表章区分の統合（小地域集計（国勢調査））</p>

秘匿処理について

(1) 秘匿処理とは

マイクロデータから世帯や個人の秘密の情報を知るということは、調査対象である調査単位(世帯や個人)とマイクロデータの対応関係を特定し、特定されたマイクロデータから調査単位の秘密に属する事項を知るということを意味する。どの調査事項が、秘密の情報に当たるかは一概には決めることができないし、時代とともに変化し、普遍的ではないと思われるので、秘匿処理とは、基本的には、調査単位とマイクロデータの対応関係を特定されないようにするというものである。

(2) 対応関係

提供するマイクロデータには、氏名、住所などの直接的に世帯や個人が特定できる情報は付与されていないので、調査単位とマイクロデータの対応関係は、性別や年齢などの属性(識別情報)が同じかどうかで判断することになる。

全国の全調査単位のマイクロデータが提供されていて、かつ、全調査単位について識別情報が分かる場合、識別情報が一致する調査単位とマイクロデータがそれぞれ1つしかない場合には同じ世帯や個人と判断でき、それぞれ複数ある場合はそのうちのいずれかと判断できる。実際のマイクロデータの提供の場合、一部の調査単位のマイクロデータが提供されていて、かつ、一部の調査単位の識別情報がわかるに過ぎず、このような状況では、対応関係を特定するのは現実的ではないと考えられる。

(3) 特定の可能性

特定の可能性を考えると、地域範囲が狭い場合には、調査対象が絞り込まれるので、識別情報を収集することが容易になり、マイクロデータの地域情報が詳細であれば、特定の可能性が高くなる。また、調査を受けていることが知られていると、その調査単位のマイクロデータに必ず存在することが分かるため、対応関係を特定される可能性が高まる。しかし、調査対象のリストは厳格に管理されており、外部の者が調査を受けている調査単位を知る可能性は低く、調査時から数年が経過すれば外部の者が知ることは不可能と言える。

しかし、特殊なデータのとくに、特定の可能性は高くなる。例えば、100歳以上の高齢者がいる世帯や世帯員が10人いるというような世帯の数は少ないので、母集団のある個別の世帯に対応するデータ数が少なくなり、そのどれに当たるか決定するのが比較的容易になる。また、複数の属性の特殊な組合せも特定の可能性が高くなる。これに対し、標準的な対象の場合には同じ条件のデータが多数出現することになるので、特定の可能性は比較的低いものととどまる。

(4) 識別情報

調査対象である調査単位とマイクロデータの対応関係を特定しようとするときに用いる識別情報とは、提供するマイクロデータに含まれていて、かつ、統計調査以外からも知ることができる情報である

個人又は世帯を対象とした統計の場合、比較的容易に入手できる識別情報としては、外観からでも把握できるような基本的な属性が考えられ、例えば、県、市町村などの地域情報や、世帯員数、世帯員の性別、住宅の大きさなどが挙げられる。このほか、自宅で営業

している世帯であればその産業・職業を知ることができるし、子供の年齢は通学している学年で分かると思われる。ただし、これらの情報だけでは、一般には対応関係を特定することはできない。また、これらの情報の収集は比較的簡単ではあるが、多数の調査単位について情報を収集しようとするれば大きな作業量を必要とする。

実際の問題としては、時間が経つとともに識別情報を正確に知ることは難しくなる。提供されるマイクロデータは数年前の調査の結果であり、そのときに個々の調査対象がどのような属性を有していたか知ることは、たとえ世帯の基本的な属性であっても難しい。既存のリストのようなものの場合も、そのリストとマイクロデータの時点が一致していないと対応関係の特定には多くの誤りが生じることになる。

(5) 特定の試み

秘匿処理の方法を決めるときには、現実にとどのような危険があるかについても考えておく必要がある。最近、個人情報の流出がよく問題となるが、そのような例では、住所（メールのアドレス等も含む。）氏名などが流出しており、それは、商業目的などにそのまま利用できる。しかし、統計情報の場合、住所、氏名が流出することはあり得ない。また、前述のとおり、特殊な対象の場合には特定の可能性が比較的高くなるが、多くの標準的な対象の場合には特定の可能性は比較的低いものにとどまる。一部の対象についてだけ特定できたとしても、商業目的での利用価値は少ないであろう。したがって、対象を特定しようとするような試みが、最近問題になっているような商業目的で行われる可能性は低いものと考えられる。そもそも、数年前の統計情報では利用する価値もないであろう。

しかし、もし対象を特定するような試みが実際に行われたら、それはマイクロデータ提供の危険性、ひいては統計調査の危険性を指摘するものとして利用されてしまうであろう。ところが、絶対的な匿名性を担保しようとする、ドイツでの経験のように提供できる情報が極めて限られてしまう。したがって、この問題は秘匿処理だけで対策を考えるべきものではなく、そのような試みを行うこと自体を制限しておくことが必要となる。このため、データを提供するときには、利用目的を限定し、データの管理を適正に行わせることを義務付けておかななくてはならない。

注：ドイツは、1980年の連邦統計法で「絶対的な匿名化」条項によるマイクロデータの提供を行ってきたが、多くの情報が失われることになり、科学研究の要求に応じられず、ほとんど利用されなかった。そのため、1987年の連邦統計法ではマイクロデータが莫大な時間や経費をかけない限り識別できないという「事実上の匿名性」の概念に法規定を改正している。

匿名化の基準（目安）

1 地理的情報について

- (1) 地理的情報としては、地域内に最小でも人口 50 万人以上いなければならない。
- (2) 直接的な地理的情報以外で、地理的情報が明らかになる項目（例えば、サンプリング情報など）についても、上記(1)の最小人口 50 万人の基準に適合させなければならない。
- (3) 地域分析用として、人口 50 万人未満の地理的情報を提供するような匿名データを作成する場合には、他の識別情報などの匿名化の程度を高めなければならない。
- (4) 入手可能な外部情報により、ある特定の種類の施設であることが明らかになるようなことがないようにしなければならない。

2 個人・世帯の識別情報について

- (1) 氏名、住所など個人又は世帯を直接的に識別できる情報は削除されなければならない。
- (2) 間接的に個人又は世帯を識別できる情報、例えば年齢、世帯人員、居住室数などの情報については、年齢の高い個人、世帯員数が多い世帯、居住室数の多い住宅など特定される可能性が高い場合、トップコーディング、グルーピングまたは削除を施す必要がある。トップコーディングにおいては、母集団（個人又は世帯）全体の 0.5% を目安にすることが望ましい。
- (3) 少数の特定の集団を対象とする場合、トップコーディングの基準を 3 ~ 5 % にすることを考慮すべきである。
- (4) トップコーディングするデータ項目については、その情報（平均値や中央値など）を明らかにすることが望ましい。
- (5) 世帯単位のデータを提供する場合、調査単位が特定されないことがないように、必要があれば、匿名化を考慮する必要がある。

3 誤差（ノイズ）

- (1) ミクロデータに誤差を加えることによって、調査データと外部情報との対応関係を特定する可能性を低めることができる。他に適当な匿名化の技法がない場合には、研究・分析上の有用性を損なわない範囲で誤差を付加することを考慮すべきである。
- (2) 誤差を加える方法としては、乱数による誤差の付加（random noise）、調査単位間の調査情報の交換（swapping）、ブランク（blank）への置換え又は補定（imputation）がある。

4 リサンプリング

ミクロデータを全て提供する場合は、その一部を提供する場合に比べて、調査単位の特定の可能性が高くなる。例えば、ある人が調査を受けたことがわかっている場合には、ミクロデータの中に必ずその人のデータがあるはずとの前提で探すことができる。したがって、必要に応じて、ミクロデータの全てではなく、一部のデータだけを提供することを考慮すべきである。

5 外部ファイルとのマッチングの可能性

- (1) ミクロデータと外部の既存ファイルのデータを突き合わせるにより調査単位が識別されるような可能性があれば、それを回避するための措置をとらなければならない。
- (2) 調査のための標本フレームが、国勢調査の母集団情報以外の情報によって提供されている場合には、調査データと標本フレームの元の情報とを一致させることが可能となるおそれがあるので、事前に回避する措置をとらなければならない。

6 その他の問題

- (1) データの一連番号、データの並び順によって、およその地域範囲が推測されるおそれがあるので、削除、付替え又は並べ替えをするべきである。
- (2) サンプリングに関する情報によっては、地理的情報以外に特定の地域や集団であることが明らかになるおそれがあるので、そのような情報は削除すべきである。
- (3) 秘密の情報のうち秘匿の必要性の高い調査項目については、その調査項目自体についてグルーピング、削除等の匿名化を施す必要がある。
- (4) 時間の経過とともに、調査データを外部情報と照合することは困難になる。提供時期は調査時点から最低限2年間以上は離すべきである。

匿名データのチェックリスト（案）
就業構造基本調査を例として

匿名データを作成する統計データの名称および年次

就業構造基本調査（平成 4 年、9 年、14 年）

1 地理的情報

- (1) 提供するファイルにはどのレベルの地理的情報が含まれていますか。匿名化のために地理的情報を加工していますか。

全国を 6 地域に区分した地域ブロック。
全国 47 都道府県を 6 ブロックに集約しています。
ブロックの構成、人口、世帯数は別添 1（省略）を参照。

- (2) 直接的な地理的情報以外に地理的情報が明らかになるような情報がありますか。

特にありません。

- (3) 地域分析用に詳細な地理的情報を提供していますか。

特に地域分析用のファイルは作成していません。

- (4) ある特定の種類の施設であることが明らかになることはありますか。

特にそのようなことはありません。

2 世帯の識別情報

- (1) 世帯の識別情報として考えられるデータ項目を挙げてください。

世帯符号、世帯人員

- (2) それぞれの識別情報について、どのような匿名化措置をとっていますか。

世帯符号について、オリジナルの符号は削除し、新たに世帯単位に一連番号を付与しています。
世帯人員が 9 人以上の世帯は削除しています。

- (3) 匿名化措置を施した場合には、その情報を明示してください。

世帯人員 9 人以上の世帯は母集団全体の約 0.07% を占めています。
世帯人員の分布は別添 2 (省略) を参照。

- (4) 世帯単位のデータを提供することに対応して特別な匿名化措置を施していますか。

特別な匿名化措置は施していません。

3 個人の識別情報

- (1) 個人の識別情報として考えられるデータ項目を挙げてください。

性別、年齢

- (2) それぞれの識別情報について、どのような匿名化措置をとっていますか。

年齢が 80 歳以上のデータについては、すべて 80 歳としています。

- (3) 匿名化措置を施した場合には、その情報を明示してください。

年齢が 80 歳以上の人は、母集団全体の約 5 % を占めています。
年齢分布は別添 3 (省略) を参照。

4 誤差 (ノイズ)

匿名化措置として、誤差を付加する方法を採っていますか。誤差を付加する方法を採っている場合には、その方法を記載してください。

誤差を加える方法は採用していません。

5 リサンプリング

匿名化措置として、リサンプリングをしていますか。リサンプリングをしている場合には、その抽出方法と抽出率を記載してください。

リサンプリングを行っている。
抽出方法は単純任意抽出法を採用し、抽出率は 80% です。

6 外部ファイル

- (1) ミクロデータを特定できる可能性のある外部ファイルは存在しますか。

そのような外部ファイルは存在しません。

- (2) 母集団情報として利用している情報は何かですか。

母集団情報として利用しているのは国勢調査の調査区情報です。
調査区内の世帯名簿は調査の一環として作成し、その世帯名簿は調査関係者以外見ることはできません。

7 その他

- (1) データの一連番号、データの並び順について、何らかの匿名化措置を施していますか。

オリジナルのデータ一連番号は削除しています。
データの並び順は、世帯単位に、乱数によりランダムな並びにしています。
ランダムな並びにしてから、データの一連番号を付与しています。

- (2) サンプリング情報によって、地理的情報以外に特定の地域や集団であることが明らかになる可能性はありますか。

そのような情報はありませぬ。

- (3) 秘密の情報のうち、特に秘匿する必要性の高い調査項目がありますか。ある場合には、どのような匿名化措置をとっていますか。

特に秘匿する必要性の高い調査項目はありません。

- (4) 提供時期と調査時点とはどの程度の期間が開いていますか。

調査による結果がすべて公表されてから、匿名データを提供しています。したがって、最短の期間でも調査時点から2年以上は開いています。

- (5) そのほか、データを匿名化するに当たり、措置していることがありますか。

特にありません。

匿名データのチェックリスト（案）
全国消費実態調査を例として

匿名データを作成する統計データの名称および年次

全国消費実態調査（平成元年、6年、11年、16年）

1 地理的情報

(1) 提供するファイルにはどのレベルの地理的情報が含まれていますか。匿名化のために地理的情報を加工していますか。

全国を6地域に区分した地域ブロック。
全国47都道府県を6ブロックに集約しています。
ブロックの構成、人口、世帯数は別添1（省略）を参照。

(2) 直接的な地理的情報以外に地理的情報が明らかになるような情報がありますか。

標本データを母集団に復元するための乗率は、都市階級別にそれぞれ固有の値になっているために、地理的情報と組み合わせると、市区町村レベルまで判明するおそれがあります。そのため、匿名化措置として、乗率を階級別に区分し、階級別のその平均値を乗率としています。
乗率の階級、平均値等は別添2（省略）を参照。

(3) 地域分析用に詳細な地理的情報を提供していますか。

特に地域分析用のファイルは作成していません。

(4) ある特定の種類の施設であることが明らかになることはありますか。

特にそのようなことはありません。

2 世帯の識別情報

(1) 世帯の識別情報として考えられるデータ項目を挙げてください。

世帯符号、世帯人員

- (2) それぞれの識別情報について、どのような匿名化措置をとっていますか。

世帯符号について、オリジナルの符号は削除し、新たに世帯単位に一連番号を付与しています。
世帯人員が9人以上の世帯は削除しています。

- (3) 匿名化措置を施した場合には、その情報を明示してください。

世帯人員9人以上の世帯は母集団全体の約0.07%を占めています。
世帯人員分布は別添3（省略）を参照。

- (4) 世帯単位のデータを提供することに対応して特別な匿名化措置を施していますか。

特別な匿名化措置は施していません。

3 個人の識別情報

- (1) 個人の識別情報として考えられるデータ項目を挙げてください。

性別、年齢

- (2) それぞれの識別情報について、どのような匿名化措置をとっていますか。

年齢が80歳以上のデータについては、すべて80歳としています。

- (3) 匿名化措置を施した場合には、その情報を明示してください。

年齢が80歳以上の人は、母集団全体の約5%を占めています。
年齢分布は別添4（省略）を参照。

4 誤差（ノイズ）

匿名化措置として、誤差を付加する方法を採っていますか。誤差を付加する方法を採っている場合には、その方法を記載してください。

誤差を加える方法は採用していません。

5 リサンプリング

匿名化措置として、リサンプリングをしていますか。リサンプリングをしている場合には、その抽出方法と抽出率を記載してください。

リサンプリングを行っています。
抽出方法は乗率階級別に標本数を比例配分し、乗率階級内は乗率を考慮した確率比例抽出法を採用しています。抽出率は80%です。

6 外部ファイル

(1) ミクロデータを特定できる可能性のある外部ファイルは存在しますか。

そのような外部ファイルは存在しません。

(2) 母集団情報として利用している情報は何かですか。

母集団情報として利用しているのは国勢調査の調査区情報です。
調査区内の世帯名簿は調査の一環として作成し、その世帯名簿は調査関係者以外見ることとはできません。

7 その他

(1) データの一連番号、データの並び順について、何らかの匿名化措置を施していますか。

オリジナルのデータ一連番号は削除しています。
データの並び順は、世帯単位に、乱数によりランダムな並びにしています。
ランダムな並びにしてから、データの一連番号を付与しています。

(2) サンプリング情報によって、地理的情報以外に特定の地域や集団であることが明らかになる可能性はありますか。

そのような情報はありませぬ。

(3) 秘密の情報のうち、特に秘匿する必要性の高い調査項目がありますか。ある場合には、どのような匿名化措置をとっていますか。

秘密の情報のうち、年間収入について秘匿の必要性を検討したが、年間収入から調査単位が特定される可能性は低いとして、匿名化措置は特に施していません。
また、年間収入は回帰分析などで説明変数としてよく利用され、ジニ係数の計算のため

にも実数でないと困ることから、利用の面も考慮してそのまま提供しています。

- (4) 提供時期と調査時点とはどの程度の期間が開いていますか。

調査による結果がすべて公表されてから、匿名データを提供しています。したがって、最短の期間でも調査時点から2年以上は開いています。

- (5) そのほか、データを匿名化するに当たり、措置していることがありますか。

特にありません。