

統計データの二次利用促進に関する研究会 報告書

平成 20 年 10 月
統計データの二次利用促進に関する研究会

目次

はじめに	1
基本的な考え方	1
二次利用の運用手続面について	1
1 利用目的の範囲	1
ア. 「公益性」を求める理由	1
(1) オーダーメイド集計	1
(2) 匿名データの提供	2
イ. 「学術研究の発展に資すると認める場合」	2
(1) 大学・研究機関における場合	2
(2) 大学・研究機関以外における場合	2
ウ. 「その他総務省令で定める場合」	3
2 利用の申出	3
ア. 利用申出書	3
(1) オーダーメイド集計の場合(統計法第34条の総務省令で定める手続き)	3
(2) 匿名データの提供の場合(統計法第36条の総務省令で定める手続き)	3
イ. 誓約書の添付	4
ウ. 利用申出者	4
エ. 利用申出の受付窓口	5
3 利用目的の審査	6
ア. 利用目的の審査内容	6
(1) 「公益性を示す書類」	6
(2) 審査内容	6
(3) 審査の実施主体	7
4 審査結果に対して不服があった場合の対応方法	7
5 集計結果表の著作権に関する留意事項	8
6 その他運用面での留意事項	8
(1) 提供窓口について	8
(2) 提供媒体・提供方法について	8
(3) 提供された匿名データの利用中の管理	9
(4) 提供されたデータの利用終了後の措置	9
(5) 提供された集計結果表や匿名データの利用範囲と成果公表方法につい て	9

(6) 匿名データを作成する統計調査について	9
(7) 二次利用の円滑な利用のための環境整備について	9
(8) 不適切利用に対するペナルティについて	10
(9) 国外の研究者に対する提供について	10
(10) レプリカデータの作成・提供について	10
(11) オンサイト利用の検討について	10

二次利用の技術的な課題について	11
1 オーダーメイド集計における集計結果表の秘匿処理	11
2 匿名データの作成	11
(1) 匿名データの作成対象	11
(2) 匿名処理の方法	11
3 秘匿処理の審査	11
4 秘匿処理の実施機関	12
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	13
別紙2 秘匿処理について	14
別紙3 匿名化の基準(目安)	18
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	20
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	24

参考資料

資料1 「統計データの二次利用に関する研究会」の開催について	1
資料2 「統計データの二次利用に関する研究会」開催実績	3
資料3 統計法における統計データの二次利用の関係条文	4
資料4 政府統計マイクロデータの試行的提供の概要	8
資料5 東京大学SSJDA(Social Science Japan Data Archive)	21
資料6 諸外国の統計データの二次的利用の状況	29
資料7 - 1 米国における秘匿処理の審査について	39
資料7 - 2 Disclosure Review Boardsの構成	40
資料7 - 3 データ開示における潜在的な露見可能性に関するチェックリスト	41

はじめに

当研究会は、統計法（平成 19 年法律第 53 号）において、委託による統計の作成等（第 34 条）及び匿名データの作成・提供（第 35 条、第 36 条）を行うことができるとの規定が設けられたことにより、これら統計データの二次的な利用（以下「二次利用」という。）を促進する観点から関連の諸課題について検討するため、総務省政策統括官（統計基準担当）の求めにより開催されたものである。研究会は平成 19 年 10 月 22 日から平成 20 年 6 月 16 日まで計 7 回にわたり、二次利用の運用手続きに関する課題、集計結果表や個別データの秘匿処理に関する技術的な課題を中心に検討を重ね、本報告書を取りまとめるに至った。

基本的な考え方

平成 21 年春の統計法の全部施行に伴い、各府省における二次利用に係る業務が行政活動の一環として適切に実施されることが必要となる。

このため、各府省では、二次利用に係る利用申出手続きや利用目的の審査等の運用に係る事務作業を府省間で整合性を確保しながら適切に実施することが必要である。また、二次利用に係る処理作業を実施するためには、国民の共有財産である行政資源を費やす必要があるとともに、特に匿名データについては、秘匿処理が施されているものの、個別客体の調査結果情報が提供されることから、調査対象者の統計調査に対する信頼を損なわないようにすることが重要である。

このような観点を踏まえつつ、当研究会では、各府省が整合性を確保しながら適切に二次利用の関係業務を実施していくための運用面、技術面における諸課題について検討し、以下のように整理した。

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している。

ア. 「公益性」を求める理由

統計法が、オーダーメイド集計や匿名データの提供に当たって利用目的を制限しているのは、次のような理由によるものである。

(1) オーダーメイド集計

オーダーメイド集計は、依頼者自身が調査票情報を利用することがないことから、提供する統計等に秘匿処理が確実に行われていれば安全性は高い。

しかしながらその実施に当たっては、相当程度の処理作業を伴い、国民の共有財産である行政資源を費やすこととなり、また、目的を問わずに依頼に応ずることとした場合には、調査対象者の統計調査に対する信頼を損ね、ひいては調査への協力が得られなくなること、また、統計の真実性の確保に支障をきたすおそれがあることから、一定の公益性を求めることとしたものである。

(2) 匿名データの提供

匿名データの提供は、秘匿措置が施されているとしても、集計されていない情報が提供されるものであり、目的を問わずに提供依頼に応ずるとした場合には、調査対象者の統計調査に対する信頼を損ね、ひいては調査への協力が得られなくなること、また、統計の真実性の確保に支障をきたすおそれがあることから、一定の公益性を求めることとしたものである。

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

典型的な場合として、大学や研究機関において教授、准教授、講師、助教、いわゆるポストドクター（博士研究員）、大学院生等が学術的な研究活動を行う場合が想定される。

しかしながら、一定程度の公益性が確保されるためには、このような研究によって社会に対して何らかの貢献が認められることが必要であり、学術論文等の形で研究の成果が公表され、社会に還元されることを要件とするように総務省令で規定すべきである。

(2) 大学・研究機関以外における場合

営利企業に属する者が企業活動の一環として研究を行う場合においても、それが学術的な研究を目的とするものであって、研究論文等の形で当該研究の成果が社会に還元される場合であれば、本要件に該当するものと考えられる。

また、学術研究目的に一部営利目的が含まれている場合であっても、当該研究の成果が研究論文の形で社会に公表・還元された後に、論文に掲載されたものが営利目的で利用されるように、主として学術研究目的で利用され、公表されたものが副次的に営利目的で利用される場合であれば公益性が認められると考える。

しかしながら、当該研究の成果の直接的な利用目的が、企業等の組織内部における業務上の資料として使用される場合や特定の顧客に対するレポート作成の基礎資料とされるような場合、あるいは研究論文として公表するもの以外の成果を作成し、顧客等に提供するような場合には公益性があるとは認められないと考える。

ウ.「その他総務省令で定める場合」

大学等の高等教育機関において講義・演習等の教育目的に利用するような場合を規定することが考えられる。この場合、大学の授業において学生がオーダーメイド集計による集計結果表（以下「結果集計表」という。）や提供を受けた匿名データを直接利用することが想定される。このような場合には、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要件を付すことが必要であり、使用するコンピュータについて、USB等外部記憶媒体への書き出しができないようにしておくことが必要である。

また、結果集計表や匿名データの取扱いについて、履修者に予備講義を実施する等、統計倫理に関する学習が行われていることを利用の前提とすべきである。

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、オーダーメイド集計や匿名データの提供を求める場合の申出事項は総務省令で定めることが規定されている。利用者の利便性を考慮すると利用申出書の様式は府省間で統一されることが必要であり、申出書に盛り込むことが必要と考えられる。申出事項は以下とおりである。

(1) オーダーメイド集計の場合（統計法第34条の総務省令で定める手続き）

- 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
（学術研究目的の場合：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））
（教育目的の場合：講義や演習（卒業研究の指導を含む。）の実施計画（統計倫理に関する教育の計画を含む。）使用方法、利用実績の報告方法）
- 集計等の作業委託内容（集計結果表様式を添付）
- 作成した集計結果表を使用する者の氏名及び所属
- 作成した集計結果表の提供を受ける方法

(2) 匿名データの提供の場合（統計法第36条の総務省令で定める手続き）

- 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
（学術研究目的の場合：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））
（教育目的の場合：講義や演習（卒業研究の指導を含む。）の実施計画（統

- 計倫理に関する学習の計画を含む。) 使用方法、利用実績の報告方法)
- 使用場所
 - データの管理方法 (データの保存場所・保存方法、コンピュータの使用環境 (インターネット非接続、外部記憶媒体への書き出し不可等))
 - データを取扱う者の範囲
(学術研究目的の場合: 共同研究者、外部委託する場合の委託先等)
(教育目的の場合: 指導教授、指導教授が認めた研究や学習の補助者、演習履修者等)
 - 匿名データの提供を受ける方法

イ. 誓約書の添付

利用者の義務や責任を明確に認識させる観点から、結果集計表や匿名データを取扱う者全員の誓約書を添付させる必要がある。誓約書に盛り込むべき事項の例は以下のとおり。

- 利用申出書に記載した以外の利用を行わない旨
- 匿名データを適正に管理する旨
- 秘密の保護に遺漏が無いよう厳重に注意する旨
- 不適切利用を行った場合のペナルティに合意する旨
- 利用終了後は提供された匿名データを直ちに返却する旨
- 研究成果の公表に際しては、統計法に基づいて関係の府省から結果集計表や匿名データの提供を受けたことを明記する旨
- 個体の特定に結びつくような研究・分析を行わない旨

ウ. 利用申出者

《学術研究目的の場合》

学術研究を目的としたオーダーメイド集計や匿名データの利用申出者は、自分自身が責任を持って研究を行うことができる者である。したがって、大学・研究機関における場合の利用申出者を例示すれば、以下のとおりとなる。

教授等の指示により、大学院生や学部学生が集計結果表の作成等に携わる場合の匿名データの利用申出者は、指導する教授等である。(注)

(注) この場合の大学院生や学部学生は、匿名データの提供を受けた者(教授等)から「業務の委託を受けた者その他の当該委託に係る業務を受託した者」に位置づけられ、統計法第 42 条第 2 項 (適正管理) 及び同法第 43 条第 2 項 (利用制限) に規定する義務が生じるとともに、同法第 61 条第 3 号の罰則規定が適用される。従って、これらの者が匿名データを不正な利益を図る目的で提供や盗用した場合にはその者が直接罰則の対象となり、その際の教授等の責任として、運用上、匿名データの利用取り消し等の措置の対象となることもあり得る。

このため、申出時においては、匿名データの利用者として、大学院生や学部学生を特

定することが必要となる

指導教授の指示により、結果集計表を用いて大学院生や学部学生が研究に必要な作業を行なう場合の申出者は、指導する教授等である。

共同研究として、教授等と大学院生が研究を行う場合の申出者は、その教授等と大学院生である。

教授等の指導監督下でなく、個人として大学院生や学部学生が研究を行う場合の申出者は、その大学院生や学部学生である。

《教育目的の場合》

教育を目的としたオーダーメイド集計や匿名データの利用申出者は、結果集計表や匿名データの利用者である。したがって、大学における場合の利用申出者を例示すれば、以下のとおりとなる。

指導教授が結果集計表や匿名データを用いて講義等の資料を新たに作成して配布する場合の利用申出者は指導教授である。

指導教授が匿名データをそのまま用いて講義や演習（卒業論文の作成等）を行う場合に、匿名データを大学院生や学部学生が利用する場合の利用申出者は利用者全員となる。（注）

（注） この場合の大学院生や学部学生は、「匿名データの提供をうける者」として、統計法第42条第1項第2号（適正管理）及び同法第43条第2項（利用制限）に規定する義務が生じるとともに、同法第61条第3号の罰則規定が適用されることとなり、これらの者が匿名データを不正な利益を図る目的で提供や盗用した場合にはその者が直接罰則の対象となる。また、その際の教授等の責任として、運用上、匿名データの利用取り消し等の措置の対象となることもあり得る。

指導教授が結果集計表をそのまま用いて講義や演習を行う場合に、集計結果表を大学院生や学部学生が利用する場合の利用申出者は教授等である。

なお、匿名データの提供の場合、秘匿措置が施されているものの、集計されていない情報が提供されるものであり、統計法第42条1項第2号で匿名データの提供を受けた者の適正な管理義務を規定している趣旨から見て、匿名データを提供する場合には、利用申出者の本人確認を十分に行う必要がある。

エ．利用申出の受付窓口

申出に際して、オーダーメイド集計の場合は、オーダーしたい結果表の作成が可能であるか否か等について調査実施者との個別事前打ち合わせが、匿

名データの利用の場合は、匿名データの仕様等について調査実施者への照会等が必要であることから、利用申出書の受付窓口としてはいずれの場合も、当該統計調査等の内容を熟知している所管の各府省とすることが効率的である。

オ．利用申出の時期

利用申出の受付から提供までの期間を明確にし、年間の受付時期・期間を年数回に限定する等の方法は、行政機関における効率的かつ円滑な運用が可能となり、ひいては利用者としても円滑な提供が期待されることから、調査実施者はこのような期間を明確にしたり、受付期間を限定することについて検討すべきである。

3 利用目的の審査

ア．利用目的の審査内容

利用目的を審査する場合、研究成果の公表方法が明確になっているとともに、「公益性」の有無がポイントであることから、申出時においては「公益性を示す書類」の添付を求めることが必要である。

(1) 「公益性を示す書類」

公益性を示す書類としては、たとえば次のようなものが想定される。

- 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)
- 学術研究成果の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)
- 教育目的で利用することを証明する書類
(高等教育機関に所属していることを証明する書類、高等教育機関が公にしている講義や演習内容の概要、シラバス等)
- 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

学術研究の発展に資すると認める場合の利用目的を審査するに当たっては、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であるかどうか、研究の成果が社会に還元(学術論文等の形で研究の成果が公表)されることになるか、提供したデータの管理体制や管理方法に信頼

がおけるか、 研究内容と利用される統計調査の関係が適切であるか等を審査する必要がある。

教育目的での利用を審査するに当たっては、 依頼者の属する機関が外形的にみて高等教育機関であるかどうか、 講義等の実施状況が報告されるかどうか、 講義で使用した資料や演習で作成した資料の概要が提示されるかどうか、 提供した結果集計表や匿名データの管理体制や管理方法に信頼がおけるか、 講義等の内容と利用される集計結果表や匿名データの関係が適切であるか、 利用する学生に統計倫理に関する教育が行われているか等を審査する必要がある。

(3) 審査の実施主体

利用目的の審査を行政機関が行うことは適当ではなく、第三者専門機関が審査すべきではないかとの提言もなされているが、次の理由から、オーダーメイド集計や秘匿データの提供に関する利用目的の審査は、当該統計調査の実施者が行うことが適当であると考ええる。

）利用目的の審査は、上記(2)に示すとおりの形式審査であり、研究内容の有用性や重要度を評価するものではないこと

）統計調査の調査対象者の信頼を損なわないためには、調査実施者として適正に審査した上でオーダーメイド集計を受託し、匿名データを提供していることを自ら確認しておく必要があること

）統計調査の内容や匿名データの内容について最も詳細に把握しているのは調査実施者自身であること

なお、提供状況については、統計法第 55 条に基づいて総務省がとりまとめて公表するとともに統計委員会へも報告することが想定される。ただし、利用目的の審査の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場合については、その件数を報告すべきであり、研究者の個人情報の保護にも配慮しつつ、申出において示された研究概要や応じなかった理由等を報告することについても検討が必要である。

4 審査結果に対して不服があった場合の対応方法

統計法第 34 条や第 36 条の規定では、それぞれ「一般からの委託に応じ、……統計等の作成を行うことができる」、「一般からの求めに応じ、……匿名データを提供することができる」とされていることから、行政手続法第 2 条第 2 号や行政不服審査法第 2 条第 1 項に規定する「処分」には該当しないと解される。このため、審査の際に、利用申出者には行政手続法に基づく処分手続は保障されず、また、審査結果に対して利用申出者が行政不服審査法に基づく不服申立てをすることはできないと考えられる。

しかしながら、二次利用という新たな制度を円滑かつ適切に運用する観点から、例えば、二次利用に関する政府統計全体の案内、相談、苦情等の窓口機能を設け、そこで苦情等の受付・斡旋等を行えるような仕組みを検討すべきである。

5 集計結果表の著作権に関する留意事項

オーダーメイド集計の場合、申出者の創意工夫や集計の事務処理を受託した者の創意工夫によって付加価値のついた集計結果表を作成することや、特殊な処理を行うなど、一連の事務処理の過程で、申出者や委託を受けた者に著作権が発生する可能性は否定できない。

このため、集計・分析等の業務を委託する場合には、成果物に著作権が発生するような処理を行わないことを契約上明確にしておくことが必要である。また、仮に、特別な分析やグラフの作成等、委託を受けた者の創意工夫の余地があるものを委託せざるを得ないような場合には、その成果物の著作権は調査実施者に帰属する旨を契約書において明確にしておくことが必要である。

さらに、申出者の著作権についても、当該申出による集計結果表の提供後に類似の申出が行われた場合に、先の申出者の著作権があるため応じられないというようなことが生じると運用上支障があることから、これを回避するための手段として、申出者は著作権を主張しないということを申出受理の条件に入れておくことも必要である。

6 その他運用面での留意事項

二次利用の円滑な運用を図る観点から、次のような仕組み等を検討することが必要である。

(1) 提供窓口について

申出書の提出から、結果集計表や匿名データの提供を受けるまでの期間を最短にする観点からすれば、提供窓口は、匿名データを作成する府省毎に置くことが効果的である。

なお、将来的には、諸外国に見られるように、統計調査票情報の保存と提供を一体化したデータ・アーカイブを構築することにより、利用者の利便性の向上を図るとともに、提供者側の事務処理の効率化を図る必要があると考える。

(2) 提供媒体・提供方法について

- オーダーメイド集計については、利用者が提供を受けた集計結果表を加工したり、論文等に掲載したりする場合に、電子的な形態であることが望ましいことから、CD-R等の電子的記録媒体の郵送、直接受渡またはオンラインにより提供することとすべきである。

- 匿名データについては、統計法第 42 条第 1 項第 2 号で適正な管理が規定されている趣旨から見て、集計結果表に比べてより安全性に配慮することが必要であり、当面、電子的記録媒体の書留による郵送または直接受渡しにより提供すべきである。

(3) 提供された匿名データの利用中の管理

提供された匿名データの利用に当たっては、外部のネットワークに接続していないコンピュータを用い、研究分析のために一時的にハードディスク等へ保存した場合は、研究分析終了後速やかに消去する等、管理を厳重にすべきである。

(4) 提供されたデータの利用終了後の措置

- 匿名データについては、秘匿処理が施されているとはいえ調査票情報であることから、統計法第 42 条第 1 項第 2 号で適正な管理が規定されていることから、調査対象者の信頼を損なわないためには、利用期間終了後は提供した CD - R 等が確実に返却されたことを提供側が確認できる仕組みを構築すべきである。また、利用期限内に申出した研究等が終了しない場合は、適切な延長手続きを行うことで延長が可能なるようにすべきである。
- 結果集計表についても、その利用目的を公益性のある場合に限定していることから、集計結果表の利用は、利用申出の際に提示された目的の範囲のみに限定すべきである。

(5) 提供された集計結果表や匿名データの利用範囲と成果公表方法について

結果集計表や匿名データを用いて作成する統計表等の利用は、利用申出書に記載した利用目的の範囲内に限定すべきであり、利用申出書に記載した公表方法でのみ公表するようにすべきである。

(6) 匿名データを作成する統計調査について

どのような統計調査について匿名データを作成するかは、各府省が利用者のニーズ等を勘案して判断することとなる。一般に匿名化が難しいとされている事業所系の統計調査についても、匿名化の対象とし、匿名化の努力が行われるべきである。なお、匿名化できない調査は、オーダーメイド集計で対応することを検討すべきである。

(7) 二次利用の円滑な利用のための環境整備について

利用者の利便性を向上させ、二次利用を促進する観点から、各府省は、毎年度当初に、どのような統計調査について、オーダーメイド集計や匿名データの提供を行うかが分かる情報をそのホームページから公表することが必要である。

また、関係学会と連携して、各府省は、どのような統計調査について、オーダーメイド集計や匿名データの提供を行うかを周知するようにすべきである。

さらに、関係学会に対して、結果集計表や匿名データの利用に関する倫理規定等を作成する等の協力を要請すべきである。

(8) 不適切利用に対するペナルティについて

匿名データを自己または第三者の不正な利益を図る目的で提供又は盗用した場合、統計法第 43 条第 2 項及び第 61 条第 3 号により罰せられる場合がある。また、利用申出書に記載した以外の利用を行った場合や提供された匿名データを適切に管理しなかった場合等、不適切な利用を行った者に対する対処方法として、不適切利用を行った者は、その後の一定期間、すべての府省において二次利用の申出を一切受け付けないとするようなペナルティを与えることが考えられる。この場合、オーダーメイド集計や匿名データの提供の契約書や利用申出書に添付する誓約書にペナルティの内容を具体的に記載しそのようなペナルティを受けることに利用者が同意する旨を記載するとともに、ペナルティの適用基準や判断主体等も規定しておくべきである。

(9) 国外の研究者に対する提供について

統計法における罰則規定のうち、同法第 62 条では、匿名データの提供を受けた者あるいは匿名データの取り扱いに関する業務の委託を受けた者等が図利目的で提供又は盗用した場合の罰則（同法第 61 条第 3 号）については、日本国外においてこれらの罪を犯した者に対しても適用すると規定されていることから、制度的な枠組みとしては、匿名データの提供は外国の研究者に対しても可能である。

国外の研究者に提供する場合には、各種手続のための書類や調査情報に関する膨大な書類等を翻訳する必要がある、そのための環境整備を検討することが必要である。

(10) レプリカデータの作成・提供について

レプリカデータは、統計法で規定している匿名データの一形態と考えるか、全くの擬似データと考えるかによりその取り扱いが異なるので、今後さらに定義を明確化した上でその作成・提供について検討する必要がある。

(11) オンサイト利用の検討について

オンサイト利用は、課題の明確化を図り、引続き検討すべきである。

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

オーダーメイド集計においては、客体数が少なく、集計結果表において特定の個人や事業所等が識別されるおそれがある場合には、秘密の保護の観点から、当該集計結果表に対して秘匿処理を行う必要がある。

しかしながら、集計結果表の秘匿処理の方法については、世帯対象か事業所・企業対象か、また全数調査か標本調査か、表章する項目によっても異なるため、一律の基準を設定することは困難である。

このため、「統計データの二次利用に関するガイドライン(仮称)」(以下「ガイドライン」という。)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すべきである。

2 匿名データの作成

(1) 匿名データの作成対象

一般的には、事業所・企業を対象とした調査のデータは匿名化が難しいといわれているが、調査によってデータの安全性や有用性は異なるため、匿名化が必要な項目や匿名化の程度も一律ではない。このため、事業所・企業の統計を一律に匿名化の対象外とすることは適当ではなく、匿名データの作成対象は、原則として、事業所・企業関係の統計を含めすべての統計調査とすべきである。

なお、匿名化できない調査は、オーダーメイド集計で対応することを検討すべきである。

(2) 匿名処理の方法

調査によってデータの安全性、有用性は異なるので、匿名化の方法は、調査ごとに決める必要がある。今後、各府省が行う秘匿処理の参考とするため、ガイドラインには別紙2に示す秘匿処理の技法や研究会においてモデルとして示した別紙3等を参考に匿名化の目安となるものを盛り込む必要がある。

3 秘匿処理の審査

統計法第35条第2項では、基幹統計調査に係る匿名データを作成する場合、行政機関の長はあらかじめ統計委員会の意見を聴かなければならないと規定している。

このような統計委員会における審査を効率的、効果的に実施するためには、実施する秘匿処理の内容を明確かつ簡潔に表現するためのチェックリストが必要であり、各府省間で統一化された様式で、各府省がそれぞれの所管の統

計調査ごとに作成することが必要である。

チェックリストの様式は各府省共通とすべきであるが、秘匿処理の内容は調査ごとに異なることから、各府省は所管の統計調査の結果について、ガイドラインに示される秘匿処理の技法や匿名化の目安を参考に必要な秘匿処理の内容を記載することが必要となる。

なお、匿名データのチェックリストのイメージは別紙 4 及び 5 に示すとおりである。

4 秘匿処理の実施機関

基本的には統計調査を所管する府省が秘匿処理を実施する必要がある。秘匿化という専門的な知識や技術的なスキルを必要とする処理を政府全体として効率的に実施していくためには、ある程度秘匿処理の実施機関を特定化し、秘匿化の業務を委託することも検討すべきである。

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(千世帯、万人等)</p> <p>(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ことである。

(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、特殊なデータのときに、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、現実には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ととどまる。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。

(6) 秘匿処理の技法

対応関係を特定しにくくする秘匿処理の方法としては、下記のような方法がある。

識別情報等の削除

対応関係を特定する危険性の高い識別情報である、世帯や居住地を直接的に特定できるような情報を削除する方法である。

識別情報のトップ・コーディング

対応関係を特定できる可能性が高くなる特殊な属性を、まとめる方法である。例えば、100歳以上の高齢者がいる世帯や世帯員が10人いる世帯の数は少ないので、対応関係を特定しやすくなるので、特に大きい値や小さい値を「○○以上」、「○○以下」というようにまとめる。海外では、トップ・コーディングされるのが対象全体の0.5%以上としている例などがある。

識別情報のグルーピング

特定の値をグループ分けして階級区分に変更する方法である。例えば、年齢を例にすると、22歳ではなく、21～25歳とする方法である。また、市町村コードなどの地域情報の場合は、外部の者にも把握しやすい情報であること、対応関係を調べなくてはならないデータの範囲を限定できることなどから特に注意が必要となる。海外では、人口10万人未満の地域区分は提供しないなどの基準が設けられている例などがある。

リサンプリング

マイクロデータをすべて提供するのではなく、そこから抽出した一部のマイクロデータだけを提供する方法である。この方法によれば、提供するマイクロデータが少なくなるので、対応関係を特定できる可能性を低下させることができる。

また、特定できたとの主張に対し、特定できたと考えることが適当ではないと主張する方法でもある。

マイクロデータのソート

マイクロデータの配列順を並べ替えることでランダムにし、対応関係を探り出すことができないようにする方法である。

別の概念からの秘匿処理の技法としては、マイクロデータから正確な対応関係を知ることができないようにする方法がある。具体的には、マイクロデータを加工して正しくないものにしてしまう方法である。

スワッピング

任意の2つの調査単位の間で、一部の調査事項の値を入れ替える方法である。

誤差の導入

マイクロデータの一部の調査事項(識別情報又は秘密の情報自体)に誤差を導入する方法である。

(7) 秘匿処理の方法の決定

上記のような問題があるものの、実際に海外で行われている秘匿処理の方法をみるとかなり詳細なデータをそのまま提供しているのが普通である。秘匿処理は、論理的に可

能性だけを考えると極めて厳しく行わなくてはならないことになるが、実際には、秘匿の必要性や利用面も考慮して現実的な判断の下で決定している。

そのような現実的な判断を行うために、海外では権威ある委員会などが処理の方法を最終承認する方式をとっている。我が国においても同様の手続きを踏むべきであり、一橋大学におけるマイクロデータの試行的提供では、統計局の「匿名標本データ作成・利用研究会」の承認を得ている。

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、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、マイ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、マイ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) マイ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)

—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及び年次

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。

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いふことから、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。

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) サンプル情報によって、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。

参 考 资 料

平成 19 年 10 月 2 日

「統計データの二次利用促進に関する研究会」の開催について

1 目 的

「統計データの二次利用促進に関する研究会」は、平成 21 年春（予定）の統計法全面施行に向けて「統計データの二次利用に関するガイドライン（仮称）」を策定するに当たり、利用者側からの意見等を反映させるとともに、技術的助言を得るために開催する。

2 検討事項

- (1) 統計法に基づく統計データの二次利用の運用について
- (2) 「匿名データの作成」における技術的な課題(匿名化の処理方法等)について
- (3) 「委託による統計の作成等」に係る技術的な課題(秘匿処理の方法等)について
- (4) 「統計データの二次利用に関するガイドライン(仮称)」について
- (5) その他統計データの二次利用に関する諸課題(オンサイト利用等)について

3 構 成 員

別紙のとおり

研究会の座長は、構成員の互選により定める。

座長は必要があると認めるときは、関係者を研究会に出席させ、意見を聴くことができる。

4 開催期間

平成19年10月から20年10月頃まで開催する。

5 庶 務

本研究会の庶務は、政策統括官（統計基準担当）付統計企画管理官室において処理する。

統計データの二次利用促進に関する研究会 構成員名簿
(50音順・敬称略)

(構 成 員)

げん だ 玄田	ゆう じ 有史	国立大学法人東京大学社会科学研究所教授
つばき 椿	ひろ え 広計	国立大学法人筑波大学大学院ビジネス科学研究科教授 (現・大学共同利用機関法人情報・システム研究機構統計 数理研究所データ科学系教授 リスク解析戦略研究セン ター長)
なか はら 中原	しげ き 茂樹	公立大学法人大阪市立大学大学院法学研究科准教授
ひろ まつ 廣松	たけし 毅	国立大学法人東京大学大学院総合文化研究科・教養学部教授
やまぐち 山口	こうぞう 幸三	国立大学法人一橋大学経済研究所付属社会科学統計情報 研究センター准教授

(オブザーバ)

総務省統計局統計調査部調査企画課長
内閣府経済社会総合研究所景気統計部長
財務省大臣官房総合政策課長
文部科学省生涯学習政策局調査企画課課長
厚生労働省大臣官房統計情報部審査解析室長
農林水産省大臣官房統計部統計企画課長
経済産業省統括経済産業調査官
国土交通省総合政策局情報管理部情報安全・調査課長
環境省総合環境政策局環境計画課企画調査室長
独立行政法人統計センター総務部総務課長
日本銀行調査統計局統計整備担当企画役

「統計データの二次利用促進に関する研究会」開催実績

第 1 回 平成 19 年 10 月 22 日開催

- 議題 : 1 研究会の運営について
2 統計法改正の概要等について
3 統計データの二次利用の事例について
4 研究会における検討課題・スケジュール案について
5 その他

第 2 回 平成 19 年 12 月 14 日開催

- 議題 : 1 「統計データの二次利用促進に関する研究会」の検討事項に係る対応方針案について
2 匿名データの作成事例について
3 その他

第 3 回 平成 20 年 1 月 28 日開催

- 議題 : 1 「統計データの二次利用促進に関する研究会」の検討事項に係る対応方針案について
2 経済産業省における研究会結果概要について
3 オーダーメイド集計における秘匿処理について
4 その他

第 4 回 平成 20 年 3 月 21 日開催

- 議題 : 1 統計データの二次利用の運用について
2 秘匿処理の審査について
3 その他

第 5 回 平成 20 年 4 月 18 日開催

- 議題 : 1 匿名データのチェックリスト(案)
2 その他

第 6 回 平成 20 年 5 月 16 日開催

- 議題 : 1 データの開示リスクについて
2 統計研修所におけるマイクロデータを使用した共同研究の試行について
3 オンサイト利用に関する制度的な対応について
4 研究会報告の中間取りまとめ骨子案について
5 政省令の検討状況について
6 その他

第 7 回 平成 20 年 6 月 16 日開催

- 議題 : 1 研究会報告の中間取りまとめ案について
2 その他

統計法における統計データの二次利用の関係条文

第一章 総則

(定義)

第二条 この法律において「行政機関」とは、法律の規定に基づき内閣に置かれる機関若しくは内閣の所轄の下に置かれる機関、宮内庁、内閣府設置法（平成十一年法律第八十九号）第四十九条第一項若しくは第二項に規定する機関又は国家行政組織法（昭和二十三年法律第二百十号）第三条第二項に規定する機関をいう。

(第2項～第10項 略)

11 この法律において「調査票情報」とは、統計調査によって集められた情報のうち、文書、図画又は電磁的記録（電子的方式、磁気的方式その他の知覚によっては認識することができない方式で作られた記録をいう。）に記録されているものをいう。

12 この法律において「匿名データ」とは、一般の利用に供することを目的として調査票情報を特定の個人又は法人その他の団体の識別（他の情報との照合による識別を含む。）ができないように加工したものをいう。

第三章 調査票情報等の利用及び提供

(調査票情報の二次利用)

第三十二条 行政機関の長又は届出独立行政法人等は、次に掲げる場合には、その行った統計調査に係る調査票情報を利用することができる。

- 一 統計の作成又は統計的研究（以下「統計の作成等」という。）を行う場合
- 二 統計を作成するための調査に係る名簿を作成する場合

(調査票情報の提供)

第三十三条 行政機関の長又は届出独立行政法人等は、次の各号に掲げる者が当該各号に定める行為を行う場合には、その行った統計調査に係る調査票情報を、これらの者に提供することができる。

- 一 行政機関等その他これに準ずる者として総務省令で定める者 統計の作成等又は統計を作成するための調査に係る名簿の作成
- 二 前号に掲げる者が行う統計の作成等と同等の公益性を有する統計の作成等として総務省令で定めるものを行う者 当該総務省令で定める統計の作成等

(委託による統計の作成等)

第三十四条 行政機関の長又は届出独立行政法人等は、その業務の遂行に支障のない範囲内において、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

場合には、総務省令で定めるところにより、一般からの委託に応じ、その行った統計調査に係る調査票情報を利用して、統計の作成等を行うことができる。

(匿名データの作成)

第三十五条 行政機関の長又は届出独立行政法人等は、その行った統計調査に係る調査票情報を加工して、匿名データを作成することができる。

- 2 行政機関の長は、前項の規定により基幹統計調査に係る匿名データを作成しようとするときは、あらかじめ、統計委員会の意見を聴かななければならない。

(匿名データの提供)

第三十六条 行政機関の長又は届出独立行政法人等は、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は、総務省令で定めるところにより、一般からの求めに応じ、前条第一項の規定により作成した匿名データを提供することができる。

(事務の委託)

第三十七条 行政機関の長又は届出独立行政法人等は、その行った統計調査に係る調査票情報に関し第三十四条又は前条の規定に基づき行う事務の全部を委託するときは、その業務の内容その他の事情を勘案して政令で定める独立行政法人等に委託しなければならない。

(手数料)

第三十八条 第三十四条の規定により行政機関の長に委託をする者又は第三十六条の規定により行政機関の長が作成した匿名データの提供を受ける者は、実費を勘案して政令で定める額の手数料を国(前条の規定による委託を受けた独立行政法人等(以下この条において「受託独立行政法人等」という。))が第三十四条又は第三十六条の規定に基づき行う事務の全部を行う場合にあっては、当該受託独立行政法人等に納めなければならない。

- 2 前項の規定により受託独立行政法人等に納められた手数料は、当該受託独立行政法人等の収入とする。
- 3 第三十四条の規定により届出独立行政法人等に委託をする者又は第三十六条の規定により届出独立行政法人等が作成した匿名データの提供を受ける者は、実費を勘案して、かつ、第一項の手数料の額を参酌して届出独立行政法人等が定める額の手数料を当該届出独立行政法人等に納めなければならない。
- 4 届出独立行政法人等は、前項の規定による手数料の額の定めを一般の閲覧に供しなければならない。

第三十九条 次の各号に掲げる者は、当該各号に定める情報を適正に管理するために必要な措置を講じなければならない。

- 一 行政機関の長 当該行政機関の行った統計調査に係る調査票情報、第二十七条第一項の規定により利用する基幹統計調査又は一般統計調査に係る調査票情報、事業所母集団データベースに記録されている情報、第二十九条第一項の規定により他の行政機関から提供を受けた行政記録情報及び第三十五条第一項の規定により作成した匿名データ

(第二号 略)

- 三 届出独立行政法人等 当該届出独立行政法人等の行った統計調査に係る調査票情報、事業所母集団データベースに記録されている情報及び第三十五条第一項の規定により作成した匿名データ

- 2 前項の規定は、同項各号に掲げる者から当該各号に定める情報の取扱いに関する業務の委託を受けた者その他の当該委託に係る業務を受託した者について準用する。

第四章 調査票情報等の保護

(調査票情報等の利用制限)

第四十条 行政機関の長、地方公共団体の長その他の執行機関又は届出独立行政法人等は、この法律（地方公共団体の長その他の執行機関にあっては、この法律又は当該地方公共団体の条例）に特別の定めがある場合を除き、その行った統計調査の目的以外の目的のために、当該統計調査に係る調査票情報を自ら利用し、又は提供してはならない。

- 2 第二十七条第二項の規定により総務大臣から事業所母集団データベースに記録されている情報の提供を受けた行政機関の長、地方公共団体の長その他の執行機関又は届出独立行政法人等は、同項各号に掲げる目的以外の目的のために、当該事業所母集団データベースに記録されている情報を自ら利用し、又は提供してはならない。
- 3 第二十九条第一項の規定により行政記録情報の提供を受けた行政機関の長は、当該行政記録情報を同項の規定により明示した利用目的以外の目的のために自ら利用し、又は提供してはならない。

(守秘義務)

第四十二条 次の各号に掲げる者は、当該各号に定める情報を適正に管理するために必要な措置を講じなければならない。

- 一 第三十三条の規定により調査票情報の提供を受けた者 当該調査票情報
- 二 第三十六条の規定により匿名データの提供を受けた者 当該匿名データ

- 2 前項の規定は、同項各号に掲げる者から当該各号に定める情報の取扱いに関する業務の委託を受けた者その他の当該委託に係る業務を受託した者について準用する。

(調査票情報の提供を受けた者の守秘義務等)

第四十三条 次の各号に掲げる者は、当該各号に定める業務に関して知り得た個人又は法人その他の団体の秘密を漏らしてはならない。

- 一 前条第一項第一号に掲げる者であつて、同号に定める調査票情報の取扱いに従事する者又は従事していた者 当該調査票情報を取り扱う業務

- 二 前条第一項第一号に掲げる者から同号に定める調査票情報の取扱いに関する業務の委託を受けた者その他の当該委託に係る業務に従事する者又は従事していた者 当該委託に係る業務

- 2 第三十三条の規定により調査票情報の提供を受けた者若しくは第三十六条の規定により匿名データの提供を受けた者又はこれらの者から当該調査票情報若しくは当該匿名データの取扱いに関する業務の委託を受けた者その他の当該委託に係る業務に従事する者若しくは従事していた者は、当該調査票情報又は当該匿名データをその提供を受けた目的以外の目的のために自ら利用し、又は提供してはならない。

第七章 罰則

第六十一条 次の各号のいずれかに該当する者は、五十万円以下の罰金に処する。

(第一号及び第二号 略)

- 三 第三十六条の規定により匿名データの提供を受けた者又は当該匿名データの取扱いに関する業務の委託を受けた者その他の当該委託に係る業務に従事する者若しくは従事していた者で、当該匿名データを、自己又は第三者の不正な利益を図る目的で提供し、又は盗用した者

第六十二条 第五十七条第一項第二号及び第三号、第五十八条、第五十九条並びに前条第三号の罪は、日本国外においてこれらの罪を犯した者にも適用する。

附 則

(準備行為)

第三条 改正後の統計法(以下「新法」という。)第六条の規定による作成基準の設定、新法第二十七条の規定による事業所母集団データベースの整備、新法第二十八条の規定による統計基準の設定及び新法第三十五条の規定による匿名データの作成並びにこれらに関し必要な手続その他の行為は、この法律の施行の日(以下「施行日」という。)前においても、新法の例によりすることができる。

政府統計マイクロデータの試行的提供の概要

一橋大学経済研究所附属社会科学統計情報研究センターは、総務省統計局統計調査部からの依頼に応じて、秘匿処理を施した政府統計マイクロデータを、全国の大学研究者に、学術研究のため提供する試行的システムを構築・運用している。

試行的提供は、政府統計マイクロデータ提供の在り方について検討する目的で、平成16年11月から20年9月までの4年間の予定で行っており、その概要は以下のとおりである。

1 提供する調査の範囲

- 就業構造基本調査（平成4年、9年、14年）
- 社会生活基本調査（平成3年、8年、13年）
- 全国消費実態調査（平成元年、6年、11年、16年）
- 住宅・土地統計調査（平成5年、10年、15年）

2 提供するデータの範囲

- 社会生活基本調査・・・生活時間と生活行動のいずれかしか調査されていない世帯員のデータは提供しない
- 全国消費実態調査・・・購入先、曜日別の品目分類のデータは提供しない
世帯票又は家計簿のいずれかが調査されていない世帯のデータも提供しない
- 住宅・土地統計調査・・・平成10年、15年調査の調査票乙のデータは提供しない

3 提供するデータ

提供するデータセットの数は1種類とする。

社会生活基本調査については、生活時間編と生活行動編を組み合わせた分析をできるように、リサンプリングにおいても、生活時間、生活行動ともに同じ世帯（・世帯員）のデータを提供する。

全国消費実態調査については、用途分類、品目分類、貯蓄・耐久財、年間収入、家計資産のデータを結合して、1つのレコード（世帯単位）として提供する。

4 秘匿処理

- (1) 地域区分は、就業構造基本調査、社会生活基本調査及び全国消費実態調査では、「北海道・東北」、「関東」、「北陸・東海」、「近畿」、「中国・四国」、「九州・沖縄」の6地域、住宅・土地統計調査は47都道府県とする。ただし、平成13年社会生活基本調査の調査票Bについては地域区分を付与しない。
- (2) 全データではなく、全体から約8割の世帯を無作為抽出したデータを提供する。ただし、住宅・土地統計調査は約1割の住宅又は住宅以外の建物とする。
- (3) 世帯員が9人以上の世帯は、その世帯全体を削除する。
- (4) 年齢については、80歳以上はすべて80歳に変換する。
- (5) 住宅・土地統計調査について、住宅の居住室数や畳数など、建物の階数はトップコーディングし、建物の建築面積と敷地面積は階級化する。
- (6) 全国消費実態調査と住宅・土地統計調査の乗率は、元の数値ではなく乗率の大きさ

階級別の平均値に変換する。

(7) データの配列順は、乱数を付してソートする方法で無作為化する。

5 利用の条件

- (1) 利用の目的は、学術研究に限定している。教育目的での利用は許可しない。営利企業からの委託研究については利用を許可しない。
- (2) 申請者は、私立大学も含めた全国の大学の講師以上の専任教員とする。申請者は、利用目的に係る研究協力者を共同利用者にすることができる。
- (3) 共同利用者には、大学の講師以上の専任教員のほか、大学の助教・助手、公的研究機関の研究者、大学や独立行政法人の「科学研究費補助金研究者」なども認めている。ただし、大学院生並びに営利企業又は民間（社団、財団等を含む）研究機関等の職員を共同利用者とすることは認めていない。

6 提供方法

- (1) センターで年4回原則として1月、4月、7月、10月（それぞれ募集期間は1か月間）に利用者を募集する。
- (2) 利用者はセンターのホームページから申請書をダウンロードして、必要事項を記入し、センターに郵送で申請する。
- (3) センターでまとめて目的外使用申請を行い、承認が得られたら、秘匿処理済マイクロデータ（CD）を作成する。
- (4) 利用者はセンターが開催する説明会に出席し、秘匿処理済マイクロデータ（CD）を受け取る。説明会では、試行的提供の趣旨、データの使用条件、データ利用上の注意事項などを説明する。
- (5) 利用者は使用期間終了後、パソコン内のデータの消去を行い、その消去報告、秘匿処理済マイクロデータ（CD）及び回答を記入したアンケート用紙をセンターに返送する。
センターでは返却されたCDをまとめて裁断処置する。

7 提供するデータの形式

データは、CSV形式で提供する。

符号表でデータの内容を明示し、先頭のレコードには符号表に記載している変数名を入れる。

8 データの使用条件

- (1) データの使用期間は6か月とする。
- (2) マイクロデータを取り扱う際には、ネットワークに接続していないパソコンをなるべく使用する。

9 研究論文

利用者は作成したマイクロデータによる研究論文をセンターに2部提出する。提出された研究論文については、随時、試行的提供のホームページ上で公開する。

政府統計ミクロデータの試行的提供における利用状況

平成16年11月から平成19年7月までの募集における申請件数は計66件で、利用者数は総計92名（うち共同利用者は26名）である。

	16年11月募集	17年4月募集	17年10月募集	18年4月募集
申請件数	6件	4件	12件	9件
(利用者別)				
申請者	6人	4人	12人	9人
共同利用者	1人	-	4人	5人
(調査別)				
就調	1件	2件	4件	7件
社会調	1件	-	1件	-
全消	4件	3件	8件	3件
(分析方法)				
表のみ	2件	-	3件	-
表+多変量解析	2件	3件	5件	8件
多変量解析のみ	2件	1件	4件	1件
(大学)				
国公立大学	1件	2件	5件	6件
私立大学	5件	2件	7件	3件

	18年11月募集	19年4月募集	19年7月募集	合計
申請件数	20件	5件	10件	66件
(利用者別)				
申請者	20人	5人	10人	66人
共同利用者	8人	1人	7人	26人
(調査別)				
就調	11件	-	4件	29件
社会調	2件	2件	3件	9件
全消	8件	3件	2件	31件
住調	2件	-	1件	3件
(分析方法)				
表のみ	6件	2件	4件	17件
表+多変量解析	12件	2件	5件	37件
多変量解析のみ	2件	1件	1件	12件
(大学)				
国公立大学	13件	2件	5件	34件
私立大学	7件	3件	5件	32件

(申請書様式)

秘匿処理済マイクロデータ使用申請書

一橋大学経済研究所附属社会科学統計情報研究センター長 殿

平成____年____月____日

申請者(所属・職名) _____

(氏名:自署) _____ 印

私(及び共同利用者)は、この度、一橋大学経済研究所附属社会科学統計情報研究センター(以下、「センター」という。)を通して秘匿処理済マイクロデータの使用を申請します。申請が承認されデータを使用するときには、下記使用条件を厳守することを誓約します。

なお、私にあって、下記使用条件に違反があった場合、今回いただく承認が取り消され、さらに、私(及び共同利用者)が今後センターを通しての秘匿処理済データの使用を一切許可いただけなくなっても不服は申しません。また、私(及び共同利用者)の氏名と違反の内容をセンターのホームページ等に掲載されても結構です。

秘匿処理済マイクロデータの使用条件

- 1 統計法の規定を遵守し、使用する秘匿処理済マイクロデータの管理は厳正に行うこと。
 - ・提供されたCD等は、使用期間中、鍵のかかる安全なところに保管すること。
 - ・提供されたCD等は、申請した使用場所以外では使用しないこと。
 - ・提供されたデータは本申請書に署名した者だけが使用し、第三者に再提供若しくは使用させないこと。
 - ・提供されたデータは本申請書に記載した集計目的だけに使用し、他の目的では使用しないこと。
 - ・パソコン等へのマイクロデータの蓄積は、一時的な場合を除いて行わないこと。
- 2 使用期間終了後、提供されたデータを直ちにセンターに返却すること。また、パソコン内のコピー等をすべて速やかに消去し、消去した旨の報告書をセンターに提出すること。
- 3 調査に関する質問等はセンターに対して行うこと。
- 4 分析結果を発表する際には、センターを経由して秘匿処理済マイクロデータの提供を受けたことを明記すること。
- 5 作成した論文等を2部センターに提出すること。
- 6 その他センターの指示に従うこと。
- 7 提供された秘匿処理済マイクロデータの使用により何らかの不利益を被ったとしても、総務省統計局統計調査部及びセンターの責任は一切問わないこと。

私は、この度、当学部(又は研究科、研究所、センター等)の専任教員である申請者が、貴センターに秘匿処理済マイクロデータの使用を申請することを承認します。

平成____年____月____日

所属長(所属・職名) _____ (氏名) _____ 公印

申請事項

1 指定統計調査の 名称及び年次	
2 調査票の使用目的	<p>研究計画の名称（20字以内にまとめてください。）</p> <p>_____</p> <p>結果の発表の仕方 報告書等 論文 研究会等で発表</p> <p>研究目的、研究方法等を以下に簡潔に記載してください。</p>
3 集計様式	別添のとおり。
4 研究資金の出所	<p>資金なし 資金交付を受けた（出所・名称を記載してください。）</p> <p>（出所：_____ 名称：_____）</p>
5 共同利用者 （氏名は自署）	<p>所属・職名：_____</p> <p>研究者番号：_____ 氏名：_____ 印</p> <p>所属・職名：_____</p> <p>研究者番号：_____ 氏名：_____ 印</p> <p>所属・職名：_____</p> <p>研究者番号：_____ 氏名：_____ 印</p>
6 使用場所	
7 申請者連絡先	<p>メール ：_____</p> <p>電話番号：_____</p>

注) この用紙に入りきらないときは、適当な用紙を添付して記入してください。

(申請書記入要領)

申請書記入要領

申請書の書き方で分からないことがあれば、本センターにメールで質問してください。

申請するときの注意

使用条件と使用条件に違反したときのペナルティーを確認した上で申請してください。申請書の記載に事実でないことがあった場合、使用条件に違反したものとみなすことがあります。

申請者及び共同利用者の資格には十分に注意してください。なお、利用者全員の署名(自署)捺印と、所属長(学部長、研究科長、研究所長、センター長等)の公印の押印が必要です。

1 指定統計調査の名称と年次

秘匿処理済マイクロデータの調査名と使用する年次を記載してください。使用するファイルの種類が分かるように、全国消費実態調査の場合は二人以上の一般世帯(又は普通世帯)と単身世帯の別を、社会生活基本調査の場合は生活時間と生活行動の別を記載してください。

なお、住宅・土地統計調査については、平成5年の調査名は住宅統計調査と記載してください。

2 調査票の使用目的

研究計画の名称を20字以内で記載し、予定している結果の発表の仕方すべてをまるで囲んでください。発表を予定していない場合、使用は認めません。さらに、提供されたデータでどのような研究を行うかが分かるように研究目的と研究方法を簡潔に説明してください。

3 集計様式

集計様式の書き方を参考に明確に記載してください。明確に記載されていない場合、使用を許可しないことがあります。また、本センターから修正を求めることがあります。

4 研究資金の出所

特に研究資金がない場合は、「資金なし」をまるで囲んでください。この項目は、営利企業からの委託研究でないことを確認するためのものですので、研究資金の提供者と資金の名称が分かるように明確に記入してください。大学から交付

された研究費の場合、「大学からの研究費」と書いてください。なお、営利企業からの研究費であっても、純粋な学術奨励のための研究補助金の場合は使用を認めます。

5 共同利用者

共同利用者の所属・職名と氏名を記入してください。氏名は自署であることを要します。「科学研究費補助金研究者」に該当する場合は、研究者番号も記入してください。

6 使用場所

集計を行う場所を、例えば、「 研究室」というように記入してください。

7 申請者連絡先

本センターからの問い合わせに使用させていただきます。申請者ご自身が所属する大学のメールアドレス及び研究室の電話番号にしてください。なお、センターからの連絡は原則としてこのメールアドレスに送りますので、使用期間中はいつも確認するようにしてください。

所属長の承認

申請に当たっては、所属長の承認が必要です。所属長の所属・職名、氏名を記載の上、公印を押印してください。

平成 年 月 日

総務大臣 殿

大学 学部 職名

氏 名 印

誓 約 書

「平成 年 調査」調査票の統計法第 15 条第 2 項の規定に基づき、目的外使用に当たっては、統計法第 14 条の規定を遵守し、調査票を適正に管理するとともに、秘密漏洩のないよう厳重に注意します。

見本

共同利用者がいる場合は、一人一人別に作成

誓約書を作成した日

平成18年11月27日

総務大臣 殿

一橋大学経済研究所附属社会科学統計情報研究センター 助教授

所属・職名

山口 幸三

自署

捺印

印

誓約書

調査年次、調査名

「平成元年、6年、11、16年全国消費実態調査」調査票の統計法第15条第2項の規定に基づく目的外使用に当たっては、統計法第14条の規定を遵守し、調査票を適正に管理するとともに、秘密漏洩のないよう厳重に注意します。

符 号 表

別紙4

作成日	2006/7/25	訂正日	2007/7/27	調査名	平成16年全国消費実態調査
調査年次	2004年(平成16年)		データ名	Zenshou2004F(二人以上の一般世帯),Zenshou2004T(単身世帯)	
備考					

	項目名	変数名	符号	備考
1	調査ID	SurveyID	Zenshou	
2	調査年	SurveyYear	2004	
3	センター使用欄	Centre		
4	レコード連番号	RecNo	0000001~	
集計区分別データの有無				
5	家計収支編(用途分類・品目分類)	KakeiDataUmu	1:あり 0:なし	
6	貯蓄編	ChochikuDataUmu		
7	耐久財編	TaikyuuDataUmu		
8	家計資産編	ShisanDataUmu		
9	地域符号	District	1:北海道・東北 2:関東 3:北陸・東海 4:近畿 5:中国・四国 6:九州・沖縄	
10	世帯符号	SetaiCode	0000001~	
11	乗率	Weight	nnnnn.n(ブランク補足,小数点以下1桁まで表章)	
12	一般・単身の別	SetaiBetsu	F:二人以上の世帯 T:単身世帯	
調査票の有無				
13	家計9月	Kakeibo_Sep	1:有り :無し	
14	家計10月	Kakeibo_Oct		
15	家計11月	Kakeibo_Nov		
16	年収票の有無	Nenshuuhyou	1:有り :無し	
17	年収票不詳	NenshuuFushou	V:年間収入不詳有り :年間収入不詳無し,年収票無し	
18	年収票貯蓄	ChochikuFushou	V:貯蓄に不詳あり :貯蓄に不詳無し,年収票無し	
19	耐久財票の有無	Taikyuuzaihyou	1:有り :無し	
20	調査月数	ChousaTsukisuu	1~3:1~3か月	
21	世帯区分	SetaiKubun	1:勤労 2:勤労及び無職以外 3:無職	
22	世帯人員	SetaiJinin	1~ 8:1~8人	
23	就業人員	ShuugyouJinin	0~ 8:0~8人	
家計を主に支える人				
24	世帯員以外の家族	世帯主との続柄	FuzaiTsunami	2:配偶者 3:子(未婚) 4:子(夫婦であるもの) 5:父母 6:兄弟(未婚) 7:兄弟(夫婦であるもの) 8:孫(未婚) 9:孫(夫婦であるもの) 0:祖父母 X:2~0以外の親族 :3か月以上の不在者なし
25	世帯員以外の家族	不在理由	FuzaiRiyuu	1:単身赴任 2:出稼ぎ 3:入院 4:その他 :3か月以上の不在者なし
家計を主に支える人以外の不在理由別人数				
26	世帯員以外の家族	入院	FuzaiNyuuin	1~9:1~9人 :3か月の不在者なし
27	世帯員以外の家族	学業	FuzaiGaku	
28	世帯員以外の家族	その他	FuzaiTa	
29	現住居等に関する事項	子の住んでいる場所	KoDoukyo	1:同居(生計同一) 2:同居(生計別) 3:同敷地内 4:近く(徒歩5分程度) 5:片道1時間以内 6:片道1時間以上 7:子はいない
30	現住居等に関する事項	単身世帯の形態	Tanshin	1:単身赴任 2:出稼ぎ 3:その他 :一般世帯
31	現住居等に関する事項	住宅の所有関係	Shoyuu	1:持ち家(世帯員名義) 2:持ち家(世帯員以外の名義) 3:民営賃貸住宅(設備専用) 4:民営賃貸住宅(設備共用) 5:農市区町村営賃貸住宅 6:公社賃貸住宅 7:社宅・公務員住宅(借り上げふくむ) 8:借間 9:寮・寄宿舎
32	現住居等に関する事項	住宅の構造	Kouzou	1:木造 2:防火木造 3:鉄骨・鉄筋コンクリート 4:ブロック造 5:その他
33	現住居等に関する事項	住居の建て方	Tatekata	1:一戸建 2:長屋建 3:共同住宅(1・2階建て) 4:共同住宅(3~5階建) 5:共同住宅(6~10階建) 6:共同住宅(11階建以上) 7:その他
34	現住居等に関する事項	水洗式トイレの有無	Suisen	1:有り 2:無し
住宅面積				
35	現住居等に関する事項	延べ床面積	NobeYuka	10~99999:1.0~9999.9㎡ 下一桁が小数第一位
36	現住居等に関する事項	うち業務用面積	GyoumuyouYuka	:業務用無し 10~99999:1.0~9999.9㎡ 下一桁が小数第一位
37	現住居等に関する事項	持家の場合	Chidai	1:世帯員名義の家で無し 2:世帯員以外の名義で無し 3:有り :持ち家以外
38	現住居等に関する事項	住宅の敷地面積	Shikichi	10~99999:1.0~9999.9㎡ 下一桁が小数第一位
39	現住居等に関する事項	建築年	KenchikuYear	21~63:昭和21~63年 1~11:平成元~11年 :昭和20年以前又は持ち家以外
40	現住居等に関する事項	持ち家以外の入居年	NyuukyoYear	1~11:平成元~11年 :昭和63年以前又は持ち家
41	現住居以外の住宅について	現住居以外の住宅の有無	TanoJuutakuUmu	1:有り 2:無し :「単身赴任」または「出稼ぎ」(単身世帯)
42	現住居以外の住宅について	用途	J1_Youto	1:親族居住用 2:賃貸用 3:その他 :入力無し
43	現住居以外の住宅について	建築年	J1_KenchikuYear	21~63:昭和21~63年 1~11:平成元~11年 V:不詳 :昭和20年以前又は入力無し
44	現住居以外の住宅について	延べ床面積	J1_NobeYuka	10~99999:1.0~9999.9㎡ 下一桁が小数第一位
45	現住居以外の住宅について	構造	J1_Kouzou	1:木造 2:防火木造 3:鉄骨・鉄筋コンクリート 4:ブロック造 5:その他 :入力無し
}				
54	現住居以外の住宅について	用途	J4_Youto	1:親族居住用 2:賃貸用 3:その他 :入力無し
55	現住居以外の住宅について	建築年	J4_KenchikuYear	21~63:昭和21~63年 1~11:平成元~11年 V:不詳 :昭和20年以前又は入力無し
56	現住居以外の住宅について	延べ床面積	J4_NobeYuka	10~99999:1.0~9999.9㎡ 下一桁が小数第一位
57	現住居以外の住宅について	構造	J4_Kouzou	1:木造 2:防火木造 3:鉄骨・鉄筋コンクリート 4:その他 :入力無し

データ利用上の注意事項（全国消費実態調査）

1 分析に当たっての全般的な注意事項

- (1) 分析に当たっては、報告書で調査方法、標本設計の方法等を十分に理解するとともに、調査票様式をよく確認してください。報告書に掲載されている結果表による分析の場合にはそれほど問題ありませんが、マイクロデータを分析する場合にはこれらのことを十分に理解しておかないと、誤りをおかしてしまうおそれがあります。
- (2) この調査の調査期間は、二人以上の一般世帯の場合は9～11月、単身世帯の場合は10～11月です。したがって、ボーナス月は含まれていないなどの問題があり、年間の収支とみて分析することはできません。
- (3) 提供データには乗率が付されています。そのことを理解して、分析するようにしてください。例えば、平均値を求めるときには加重平均する必要があります。
- (4) 提供データにはトップコーディングやリサンプリングなどの秘匿処理を施してあります。このため、提供データによる集計結果を公表されている結果と完全に一致させることはできません。なお、秘匿処理の方法は「参考1」のとおりです。
- (5) 多変量解析を行うときには、外れ値を考慮するように注意してください。調査票の記入誤りなどはチェック・修正済みですが、外れ値を除くような処理は行っていません。

2 データの形式等について

- (1) データはCSV形式です。先頭のレコードに変数名を入れてあります。
- (2) 1世帯1レコードになっています。用途分類、品目分類、貯蓄・耐久財、年間収入、家計資産のデータを1レコードにまとめています。
- (3) 世帯票と家計簿（1か月分以上）の両方が提出されている世帯のデータだけを提供しています。世帯によっては他の調査票が提出されていない場合もありますが、その場合、該当項目の欄は欠損値（.）、不詳（V）、0又は空白になっています。調査票の提出状況については、「集計区分別データの有無」の欄及び「調査票の有無」欄で確認してください。
- (4) データの符号の付け方等については、符号表を参照してください。特殊な符号の付け方になっていることもありますので、必ず確認してから分析してください。
- (5) データ中の世帯属性等の欄については0補足を行うなどして桁数をそろえてありますが、用途分類等の欄については必要桁数だけにしています。
- (6) 当該マイクロデータが分析の対象であることを確認してから集計するようにしてください。必要のない欄に符号が付いていること（過剰記入）もあります。

(7) 各年次のデータ数については、「参考2」のとおりです。

3 データの内容について

- (1) 用途分類及び品目分類の数値は、調査期間中の収支金額を調査月数で除した1か月当たりの平均値にしています。提供するデータでは整数になるよう四捨五入しています。調査月数は用途分類の「調査月数」欄に入っています。なお、総務省統計局で行っている用途分類及び品目分類の集計では、一般の（普通）世帯の場合、2か月しか調査していない世帯の乗率は3分の2に、1か月しか調査していない世帯の乗率は3分の1にしています。単身世帯の場合及び集計世帯数の集計の場合も同様な考え方で集計されています。
- (2) 平成元年のデータには家計資産はありませんが、用途分類の末尾に「住宅・宅地資産額」と「耐久消費財資産額」だけが入っています。平成6年、11年及び16年のデータにおいて、耐久財・貯蓄等の調査票が未提出などの理由で資産データが作成されていない場合、家計資産額はすべて欠損値（.）にしています。
- (3) 年間収入調査票が提出されていない場合、世帯属性の年間収入の欄は0又は不詳（V）となっています。ただし、用途分類の欄には総務省統計局で推計した年間収入の値（世帯の合計年間収入）が入っています。
- (4) 乗率は、平成11年、16年の単身世帯の場合は比推定済乗率、他はすべて線形推定乗率になっています。乗率は小数点以下1桁まで表示しています。
- (5) 地域は「北海道・東北」、「関東」、「北陸・東海」、「近畿」、「中国・四国」及び「九州・沖縄」の6区分にしています。
- (6) 平成元年の普通世帯のデータでは、「世帯員情報」の欄（の多分、最後）に不在世帯員の情報が入っていますので、「行番号」の欄で確認した上で利用するようにしてください。
- (7) 「住宅の延べ床面積」や「宅地の敷地面積」の欄は、平成元年の場合は整数の値が入っていますが、平成6年、11年及び16年の場合は10倍した値（小数点以下1桁目までの値）が入っています。ただし、小数点以下1桁目が0のとき、0ではなく空白になっていることもあります。

お願い

マイクロデータの利用において疑問が生じたときには、当センターにメールでご質問ください（micro@ier.hit-u.ac.jp）。なるべく早くお答えするようにします。

また、今回の提供データの作成に当たっては、十分に注意したつもりですが、思わぬ誤りが残っているかも知れません。データを利用して疑問が生じたときには、当センターにメールでご連絡いただくようお願いいたします。

参考1 秘匿処理の方法

- (1) 地域は「北海道・東北」、「関東」、「北陸・東海」、「近畿」、「中国・四国」及び「九州・沖縄」の6区分にしています。
- (2) 全データではなく、全体から約8割の世帯を無作為抽出したデータを提供します。
- (3) 世帯員が9人以上の世帯は、その世帯全体を削除しています。
- (4) 世帯の配列順は、乱数を付してソートする方法で無作為化しています。
- (5) 年齢については、80歳以上はすべて80歳に変換しています。

参考2 データ数

年次	一般の（普通）世帯	単身世帯
平成元年	44,778	3,288
平成6年	44,803	3,772
平成11年	44,537	4,013
平成16年	44,006	4,001

東京大学 SSJDA (Social Science Japan Data Archive) について

本資料は、本研究会事務局が東京大学のホームページから関係部分をダウンロード及び要約して作成したものであり、東京大学社会科学研究所附属日本社会研究情報センターの了解を得て本研究会の資料としたもの。

1 概要

データアーカイブは、統計調査、社会調査の個票データ（個々の調査票の記入内容。マイクロデータ）を収集・保管し、その散逸を防ぐとともに、学術目的での二次的な利用のために提供する機関。

東京大学社会科学研究所附属日本社会研究情報センターは、我が国における社会科学の実証研究を支援することを目的として、SSJ データアーカイブ (Social Science Japan Data Archive) を構築、個票データの提供を 1998 年 4 月から開始。

2006 年 3 月末現在、SSJ データアーカイブでは 579 調査、772 データセットを保管し、2005 年度一年間で、利用申請件数は 335 件、提供データセット数は 2527 に上る。またこれまでに SSJ データアーカイブが提供して作成された論文や著書の数、同じく 2006 年 3 月末現在、総計で 283 点。

2 利用条件

(1) 利用対象者

大学又は研究機関の研究者、教員の指導を受けた大学院生。

民間の研究機関の者は原則として利用できないが、SSJ データアーカイブへ個票データを寄託している場合は利用可能。

(2) 利用目的

学術目的の二次分析に限られる。

(3) 一度に利用できる調査数

一度に利用申請できる調査数は原則 2 調査。

ただし、一度に 3 調査以上を利用申請する場合は、別途理由書を添付。

(4) 利用承認

利用者からの利用申請書を、データの寄託者又は SSJ データアーカイブが承認することが必要。

(5) 利用期限

原則として、データが提供されてから 1 年間。

利用期限後は、個票データを消去、CD-R を破壊。

3 利用手続の流れ(全体の流れ図は別紙 1 参照)

(1) 利用したい個票データの決定

ホームページの「収録調査の検索」ページを使って利用したい個票データを決定。

なお、個票データに関する情報は、調査毎に概要ファイルと調査票ファイルを組にして提供している。

(2) 個票データ利用申請書を提出

利用申請書（様式は別紙 2 参照。ホームページからダウンロード可）に必要事項を記載の上、生 CD - R と返信用封筒を同封の上、SSJ データアーカイブへ郵便で送付。

利用申請書を受領してから、利用承認の可否の審査。通常 1 か月前後を要する。

(3) 書込済 CD - R 等の送付

利用が承認された後、データを書込済の CD - R、利用承認書、利用申請書の写し、利用報告書様式（様式は別紙 3 参照。ホームページからダウンロード可）が利用者へ送付される。

(4) 利用者による二次分析

利用者による個票データを利用した二次分析の実施。利用期限は、データ提供日から 1 年。

(5) 利用期限終了後の措置

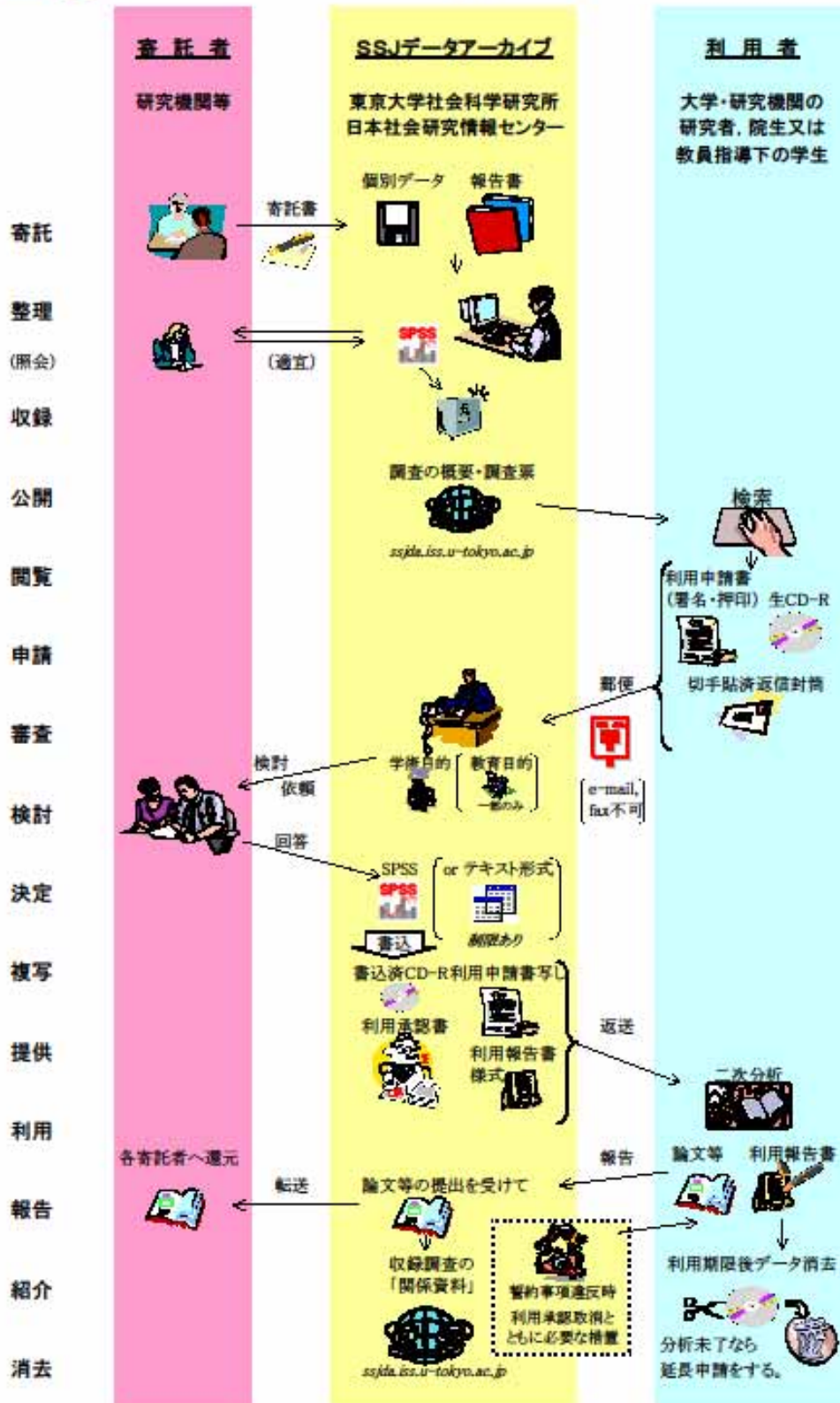
利用期限が終了した際には、個票データを消去、CD - R を破壊する。

また、利用報告書を SSJ データアーカイブへ提出するとともに、論文等を発表した場合は、それも提出する。

なお、利用期限中に二次分析が終わらない場合は、利用の延長も可能。延長する場合は、再申請が必要。

(6) 利用申請時に誓約した事項に違反した場合の措置

データ利用申請時の誓約事項に違反した場合の取扱規程(別紙 4 参照)を定め運用。



SSJDA 使用欄	申請番号	延長

個票データ利用申請書

東京大学社会科学研究所附属
日本社会研究情報センター
SSJデータアーカイブ 御中

所属（大学・学部・学科または所属機関・部課を明記）

職名又は学年
ふりがな
申請者

下記の個票データの利用について、生CD-Rおよび切手貼済返信封筒を添えて申請します（個票データは、SPSSではなくテキスト形式を希望）。

申請日 _____ 年 _____ 月 _____ 日 利用期限 データ提供日から1年
調査番号 _____ 調査名 _____ 寄託者 _____

注意： 1枚の申請書に2つ以上の寄託先を記入することはできません

二次分析のテーマ（教育目的が認められるデータを当該目的で申請する場合は、大学名・授業名）
※研究目的の方は、当該分析の具体的な計画を3頁目に示す様式に記入してください。

研究経費（該当する費目にチェックをつけること）

校費 科研費（課題番号： _____） その他の研究費 私費

（欄不足の場合：別紙記載，資料添付も可。利用延長の場合：以前の申請番号を付記し，利用報告書を添付。）

所属機関所在地

〒 _____ - _____
Tel _____ Fax _____ E-mail _____

自宅住所

〒 _____ - _____
Tel _____ Fax _____ E-mail _____

指導教員（申請者が大学院生（又は大学の学部学生）の場合）

自署 _____ 所属（大学・学部・学科を明記） _____ 職名 _____
印 _____

Tel _____ Fax _____ E-mail _____

誓約事項

- 1 提供された個票データは、学術目的での二次分析にのみ利用します。

教育目的で利用します (を付けて下さい。)

受講生も共同利用者として自署、押印が必要 (JEDSを除く)。申請時に受講者未定の場合は、
____月 ____日 まで (申請時から3ヶ月以内) に確定した共同利用者を提出します。

個別データの秘密保護を図り、個々の調査対象を特定する分析は行いません。提供された個票データは、本利用申請書に署名した者だけが利用し、第三者には再提供しません。

- 2 調査に関する照会は、原則としてSSJデータアーカイブを通して行い、寄託者に直接行いません。

- 3 二次分析の結果を発表する際には、個票データについて以下の文を付すことにより、個票データの出典を明記します。

〔二次分析〕に当たり、東京大学社会科学研究所附属日本社会研究情報センターSSJデータアーカイブから〔「〇〇〇調査」(寄託者名)〕の個票データの提供を受けました。

The data for this secondary analysis, "name of the survey, name of the depositor," was provided by the Social Science Japan Data Archive, Information Center for Social Science Research on Japan, Institute of Social Science, The University of Tokyo

- 4 個票データ利用期限までに利用申請書の内容について異動が生じた場合は、速やかにSSJデータアーカイブにその旨を届け出ます。

- 5 利用期限終了後は、個票データを消去して、利用報告書をSSJデータアーカイブに提出します。その際、論文等¹を発表していれば、利用した調査の寄託者数 + SSJデータアーカイブ分 1 の部数を同封します。

- 6 その他、SSJデータアーカイブの指示に従います。

- 7 提供された個票データ等の利用により何らかの不利益を被ったとしても、寄託者及びSSJデータアーカイブの責任は一切問いません。

利用に当たり、上記誓約事項を厳守します。

自署 印

共同利用者 (申請者以外に利用者がある場合。欄が足りない場合、別紙記載)

所属 (大学・学部・学科または所属機関・部課を明記)

職名又は学年 自署 印

所属 (大学・学部・学科または所属機関・部課を明記)

職名又は学年 自署 印

署名、押印の漏れがある場合、再提出を求められることがあります。

¹ 著書や論文、学会の大会報告書、授業等の実習報告書などの印刷されたものを指します。「論文等」が、印刷物の一部分であり全体の提出が困難な場合、該当する「論文等」を複写したのもでも可。

提出のあった論文等は、SSJデータアーカイブから各寄託者に1部ずつ転送します。また、論文等のタイトル、著者名等を、SSJDAウェブサイトにおける当該調査の「概要」ページで関連論文として紹介します。

SSJDA 使用欄	申請番号
--------------	------

二次分析の計画

<p>研究題目 (研究内容をあらわすのに適切な題目を簡潔に記入してください)</p>
<p>二次分析の内容 (研究の意義、仮説または予想される結果などについて、200字以上で記述してください)</p>
<p>利用予定の質問項目 (申請する調査データのWeb上の概要ページに含まれる主要調査項目リストを参考にしてください)</p>
<p>本研究成果の公表予定 (学会大会名、雑誌名、書籍名など、および報告・出版予定時期を記述してください)</p>

※注意事項

個票データに関する情報は、調査ごとに概要ファイルと調査票ファイルを組にしてまとめています。調査方法や質問文などについて十分に理解してください。

また、誤解に基づく分析を避けるためにも、調査報告書を必ず読んでから個別データの利用を始めてください。SSJDAでは、提供する調査の報告書を収集し、東京大学社会科学研究所図書室で閲覧できるようにしていますのでご利用ください。

なお、閲覧申請の際は調査番号を必ず確認のうえお申し込みください。

SSJDA 使用欄	論文番号	更新		転送
		申請リスト	gaiyo	

個票データ利用報告書

_____年 ____月 ____日

東京大学社会科学研究所附属
日本社会研究情報センター
SSJ データアーカイブ 御中

所属

職名又は学年

利用者

貴センターから提供を受けた個票データについて、
下記の通り利用しましたので、報告します。

記

データ提供日 _____年 ____月 ____日 (提出はデータ提供日から1年以内(他に指示がある場合を除く))

利用した個票データ

申請番号	調査番号	調査名	寄託者
_____	_____	_____	_____
_____	_____	_____	_____
_____	_____	_____	_____

1. 上記データの扱いについて (利用期限が設けられていない場合は、記入不要)

- を1つ付けて下さい。
- ア 個票データを消去した。
 - イ 個票データの利用を延長申請する (利用申請書を同封)。

2. 上記データを利用した結果について

- を1つ付けて下さい。
- ア 論文等を発表したので、同封した。
 - イ 論文等を発表予定であり、発行後速やかに SSJDA に送付する。

論文等タイトル	発表媒体名	著者名
_____	_____	_____
発行(予定)日		
_____	_____	_____
 - ア・イの場合

利用した調査の寄託者数	SSJDA	同封部数
_____	1	_____
()	+	= 計 _____

同封部数は、最低2。参照 http://ssjda.iss.u-tokyo.ac.jp/report-form.html#how_to_count
 - ウ 論文等の発表はなかった。

備考

論文等には、出典明記(誓約事項3)のページに付箋(目印となる紙片)を付けて下さい。
抜き刷りなどの場合、発表した雑誌の名前・年月などがわかるようにして下さい。



データ利用申請時の誓約事項に違反した場合の取扱規程

1. 本規程は、データ利用申請時の誓約事項に違反した場合の取扱について、これまで S S J D A で運用されてきたルールを明文化したものである。
2. 利用報告書の遅滞があった場合
現行の利用者管理システムでは、1年間の利用期限が経過したときに利用報告書の提出を督促するメールが S S J D A から利用者に対して自動的に送信される。そのメール送信後1ヶ月たっても報告書が提出されない場合、S S J D A では、再度、報告書と遅滞理由を述べた弁明書（署名捺印を要する）を郵送で提出するよう求める。この報告書および弁明書の提出日は消印の日付とする。S S J D A では、その利用者による新たなデータの利用申請を報告書および弁明書提出日から6ヶ月間受け付けない。なお、報告書と弁明書の提出が求められたにもかかわらず提出されない場合は、新たな利用申請が認められないだけでなく、当該データの寄託者にその利用者氏名を通知することとする。
3. 出典の明記（acknowledgment）がなされなかった場合
学術雑誌の場合
校正までの段階で S S J D A に連絡があった場合は、校正の際に利用者が修正するよう伝える。また、校正に間に合わなかった場合には、次号の記事として掲載するように伝える。そのうえで、利用者は弁明書（署名捺印を要する）と修正または出版社に対する修正依頼ないし掲載依頼をおこなった証拠となる文書および当該部分の複写を S S J D A に郵送で提出する。これらの提出日は消印の日付とする。ただし、いずれの場合でも、S S J D A はその利用者による新たなデータの利用申請を弁明書提出日から6ヶ月間受け付けない。

出版物の場合
校正までの段階で連絡があった場合は、学術雑誌の場合に準ずる。校正に間に合わなかった場合は、できるだけ早く当該出版物に訂正紙を添付し、二刷以降では明記するよう伝える。そのうえで、利用者は弁明書（署名捺印を要する）と修正または出版社に対する修正依頼ないし掲載依頼をおこなった証拠となる文書および当該部分の複写を S S J D A に郵送で提出する。これらの提出日は消印の日付とする。ただし、いずれの場合でも、S S J D A はその利用者による新たなデータの利用申請を弁明書提出日から6ヶ月間受け付けない。
4. 罰則が適用される範囲
利用者が学生（大学院学生を含む）である場合、利用申請書に記名・捺印されている指導教員に対しても、S S J D A から厳重な注意をおこなう。
5. 本規程の施行
本規程は、2007年4月1日より施行する。利用者に対しては、ホームページの掲載により、周知することとする。

国名	アメリカ	
	ICPSR	TheDataWeb
データアーカイブ機関等 [主な提供データの名称]	The Inter-university Consortium for Political and Social Research	U.S. Census Bureau と Centers for Disease Control のコラボレーション
機関又はデータの概要	University of Michigan 内のISR(Institute for Social Research)の一機関。 4つの運営部門(Collection Development, Collection Delivery, Educational Resources, Data Security and Preservation)を設置。 ICPSRは12人のGoverning Council(任期4年)で運営されている。	オンライン・データライブラリのネットワークで、「DataFerrett」という専用のシステムを用いてデータにアクセスする。 「DataFerrett」のセンサス局の開発担当はDSD(Demographic Surveys division)の Survey Modernization Programming Branch
運営経費の負担 (スポンサー)	会費(550以上の大学等機関)、補助金、奨励金	U.S. Census Bureau と Centers for Disease Control が 「DataFerrett」の開発には当初労働統計局(BLS:Bureau of Labor Statistics)が、後オーストラリア統計局(ABS)が資金援助していると思われる。
マイクロデータ提供開始時期 又はアーカイブ設立時期	1962年	DataWebそのものについては不明だが、「DataFerrett」の初期バージョンにあたる「Ferrett」については1997年にU.S. CensusBureauとBLSにより共同開発した。開発経費がかかるため、現在のシステムはオーストラリア統計局(ABS)と共同開発。
提供されるマイクロデータ	6300の研究調査における50万ファイルを保管。学問の分野は政治学、社会学、人口統計学、経済学、歴史、教育、老人学、刑事裁判、公衆衛生、外交政策、保健、医療、教育、保育研究、法律と薬物乱用などあらゆる領域に及び。	1990年と2000年の人口住宅センサスのPUMSのほか、ACS, AHS, BRFS, CES, CPS, NAMCS, MORT, NHANES, NHIS, NHAMCS, FHVAR, SIPP, SPDの13種類 *但しpublic use dataのみ
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	College and university Institutional Review Boards (IRBs).	センサス局内の統計開示評価委員会(Disclosure Review Board)
秘匿方法	ICPSRが寄託データの秘匿処理をすることはない。しかし、秘匿すべき情報が存在していないか、ICPSRのスタッフがチェックを行う。(ICPSRはデータ作成者に対して寄託前に調査対象の識別子をデータからすべて削除することを推奨している。)	下記の手法を用いている。 ・Removal of direct identifiers ・Setting geographical population thresholds ・Data swapping ・Global recoding ・Rounding ・Top Coding ・Age detail
利用条件・手続	データの利用の際にはICPSRに加盟している必要がある。加盟のための条件は以下の通り 1. 高等教育施設、高等教育施設を伴う団体、教育団体組織 2. ICPSRの細則に従うと認められる組織や団体	DataFerrett(Federated Electronic Research, Review, Extract, and Tabulation Tool)という専用のブラウザをダウンロードし、インストールしておく必要がある。(無料) E-mail Addressの登録(loginの際に必要) * login画面にE-mailアドレスを入力し「public use data only」というチェックボックスにチェックをしてloginする。
提供形態	HP上からFTPによるダウンロード Diskettes-3-1/2"diskette(PC or Macintosh) Zip disks-100 or 250 megabyte(PC or Macintosh) CD-ROMs-formatted to the ISO 9660 standard Tapes 8 mm cartridge tape in standard UNIX formats: tar,dd, or cpio 4 mm cartridge tape in standard UNIX formats: tar,dd, or cpio 3480 cartridge tape,38000 bits per inch(BPI),ANSI or IBM labeled 9-track round tape,6250 BPI,ANSI or IBM labeled	DataFerrettを通したオンライン上でのダウンロード DataFerrett上でのクロス集計表の作成、地図上でのデータの加工なども可能
提供フォーマット	SAS SPSS Stata	SAS SPSS Stata EXCEL タブ区切り、カンマ区切り
マイクロデータの利用料金	ICPSR加盟者の加盟資格により年会費が異なる 2004-2005年会費 Category A \$12,730 Category B \$7,960 Category S \$3,980 Category C \$2,000 Category F \$15,000 Category N \$15,000 参加加盟資格 Category A 社会科学及び関連分野の大学院を持つ大学及び研究機関(研究員60名以上) Category B 社会科学及び関連分野の大学院を持たない大学(学生数2500人以上) Category S 社会科学及び関連分野の大学院を持たない大学(学生数2500人未満) Category C 発展途上の機関 Category F Category N	無料
データの寄託元 寄託手続き	大学等の民間の研究機関(者)、政府機関 Guide to Social Science Data Preparation and Archiving (3rd edition)	U.S. Census Bureau, Bureau of Labor Statistics, Centers for Disease Control等の政府機関 -
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.icpsr.umich.edu/ http://www.icpsr.umich.edu/org/mission.html	http://dataferrett.census.gov/index.html
アーカイブスタッフ	Staff Directory ICPSRにおける管理スタッフは100名以上	センサス及びACS以外のアーカイブはDSDが所属するADDP(Associate director for Demographic Program)内で実施されていると思われるが詳細は不明。

諸外国の統計データの二次的利用の状況

国名	アメリカ	
データアーカイブ機関等 [主な提供データの名称]	センサス局 (U.S. Census Bureau)	
機関又はデータの概要	[PUMS(Public Use Microdata Sample)]	CES (the Center for Economic Studies) & RDC (the Research Data Centers)
機関又はデータの概要	人口住宅センサスから作成した匿名標本データ	CES (Center for Economic Studies)はセンサス局の一部門で、そのネットワークとなるRDC (Research Data Centers)は全米に9箇所ある。
運営経費の負担 (スポンサー)	センサス局	運営費についてはセンサス局(CES)とRDCパートナーであるacademia又は非営利団体が負担しているが、研究者にも使用料として応分の負担を求めている。(ただし、研究のスポンサーとなってもらえる機関の斡旋も行うため、標準的な料金として調整され、追加に発生するのは特別に新たなデータセットを作成したり、データのリンクページなどに伴うものだけとなる場合もある。)
マイクロデータ提供開始時期 又はアーカイブ設立時期	1963年 1960年人口住宅センサスデータの0.1%抽出データが最初	1983年にセンサス局にthe Center for Economic Statisticsを設置したのが、RDCの始まり。その後、1998年にRDCの拡大計画が実施され、現在に至る。
提供されるマイクロデータ	現在提供されているのは2000年人口住宅センサスの1%と5%のPUMS、5%PUMSは詳細な地域属性(地域区分はPublic Use Microdata Area(PUMA)でcountyレベル、最低10万人以上の単位でさらに分割。)1%PUMSは詳細な項目区分(地域区分はSuper-Public Use Microdata Areaで複数のPUMAを包含する最小40万人以上の地域。)	オンライン利用施設として、人口調査系データのみならず、DataFerrettやPUMSで公開できない経済系データや研究利用しやすく加工された43のリサーチデータを用意している。
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	統計開示評価委員会(Disclosure Review Board) (発足当初はマイクロデータ検討委員会(Micro data Review Panel)) 統計開示評価委員はセンサス局のあらゆる部門から召集された十人程度の職員で構成される。構成員の職位はアシスタントディビジョンチーフやブランチチーフ、当該調査担当者などである。	同左
秘匿方法	Data swapping(スワッピング) Top coding(トップ・コーディング) Geographic population thresholds(地理的人口境界) Age perturbation(年齢攪乱(10人以上世帯を対象)) Detail for categorical variables (is collapsed)(変数の詳細分類(の統合))	オンライン利用施設であるため、原則としてPUMではない個票データを扱えるが、センサス局職員と外部の研究者では利用できる個票データの範囲が異なる。例えば、外部の研究者には経済センサスのデータを利用する際には、調査によって得られたデータか、IRS(内国歳入庁)の税務データによって補定されたデータかわからないようになっており、企業名等の情報が削除している。
利用条件・手続	用途及び使用方法を問わない。	マイクロデータを扱うCES又はRDCの施設を利用する際には、連邦政府職員として現に雇用される必要があるため、連邦政府職員としての雇用についての宣誓とともに、Title13に関する教育訓練を受け、守秘義務履行の宣誓と宣誓供述書の署名による特別宣誓資格(SSS: Special Sworn Status)を得て、初めて可能となる。 研究利用の申請には研究内容、研究に利用するデータのほかに、センサス局にとってのメリットを記した提案書を出す必要がある。なお、利用するデータがIRSのデータが入っている経済関係データの場合、IRSの基準をクリアしなければならない。施設に入るに当たっては、研究者の背景(どこの(会社)の人か、どこかと契約して調査しているのか)など、非常に細かく調べられ、センサスパッチの着用も義務付けられている。
提供形態	HP上からFTPによるダウンロード(州別) CD等のメディアの販売。	オンラインによる提供
提供フォーマット	「dat形式」	不詳
マイクロデータの利用料金	無料(FTP方式によるダウンロード) CD等の媒体は有料	無料
データの寄託元 寄託手続き	センサス局自身	センサス局自身 研究者によるデータの持ち込みも、所定の手続を行うことで可能。
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.census.gov/main/www/access.html	http://www.ces.census.gov/index.php/ces/1.00/cms/home http://www.ces.census.gov/dialog.php/ces/1.00/ourmission Research at the Center for Economic Studies and the Research Data Centers: 2005 http://www.ces.census.gov/index.php/ces/1.00/cms/downloads
アーカイブスタッフ		上記年次報告書にP59～60にスタッフについての記述あり CENTER FOR ECONOMIC STUDIES (CES) STAFF LISTING 2005

諸外国の統計データの二次的利用の状況

国名	カナダ	
データアーカイブ機関等 [主な提供データの名称]	カナダ統計局 (Statistics Canada)	
	[PUMFs (Public Use Microdata Files)]	DLI (Data Liberation Initiative)
機関又はデータの概要	人口センサス等から作成した匿名標本データ	・発足時 Social Science Federation of Canada (SSFC)、Canadian Association of Research Libraries (CARL)、Canadian Association of Public Data Users (CAPDU)、カナダ統計局と他の政府機関の協力的なプログラムとして発足(5年間のみ) ・現在 カナダ統計局内のLibrary and Information Centre
運営経費の負担 (スポンサー)	1986年より、Canadian Association of Public Data Users (CAPDU)、とCanadian Association of Research Libraries (CARL)、そしてカナダ統計局の5つの部門がPUMFsの作成と普及に係る費用を負担	上記の組織と考えられる
マイクロデータ提供開始時期 又はアーカイブ設立時期	1973年 1971年CensusからPUMFsを作成したのが最初 (1960年代にはオーダーメイド集計で対応していた)	1996年 2月に大蔵委員会(Treasury Board)より承認を受け、同年3月に連邦政府機関(Science and Technology Strategy)となる
提供されるマイクロデータ	センサスをはじめとする22種類のデータ(2006/09/21現在) PUMFs(1%) データはおよそ4分類に大別 HEALTH SOCIAL(うちセンサスは個人・家庭・世帯に区分される) LABOUR EDUCATION	Data Base・・・14種類 Geographic Files・・・9種類 Census(PUMFs)・・・3種類 Other Micro Data Files(PUMFs)・・・7.9種類
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	・Privacy Impact Assessment(個人情報影響評価プログラム) ・Data Access and Control Service Division(アクセス・コントロール管理部門) ・Confidential and Legislation Committee(所管の4機関において秘匿の技術問題や、守秘性についての推断)	左記と同様と考えられる
秘匿方法	一般的に2つの方法を採用 Top, Bottom Cording(トップ、ボトムコーディング) Variable Categories Wider(大分類化) Removing Some of the Variables from Some Respondents(変数削除) Removing Some of the Respondents from the File(データ削除) Adding Random Noise to the Microdata(ノイズ導入) Data Swapping(スワッピング) Replacing Small Groups with Average Values(平均化) Deleting information from Some Respondents and Replacing it with Imputed Values(補定)	左記と同様と考えられる
利用条件・手続	マイクロデータ提供委員会(Microdata Release Committee)の承認を得たあとに一般に販売される。PUMFs、完全な個票の場合のいずれもライセンス合意が必要 ただし、完全な個票の利用は統計局職員扱いとなり、契約書にも明記されている 提供に関してはRDCを通じて行われる ○PUMFsライセンス契約書(MICRODATA Licence AGREEMENT) (オンライン契約書はRDCを参照)	年間ライセンス料を支払う高等教育機関のスタッフと学生 学術研究に目的は限られており、利用目的を明確にしなければならない
提供形態	CD-ROMによる購入 DLIの支援により、オンライン提供もされている	Web上でダウンロード FTP(またはファイル転送プロトコル) CD-ROMによる購入
提供フォーマット	-	-
マイクロデータの利用料金	ファイルごとに料金を徴収する(1000～3000ドル程度) 一括購入の際に割引があることもある 教育機関についてはカナダ統計局のマイクロデータファイルを50%引きで購入できる (現在の価格は不明)	規模の大きな機関(27機関)については、年間12,000ドル (the Canadian Association of Research Libraries [CARL]) その他の機関については年間3,000ドル (Canadian Association of Small University Libraries [CASUL]等の機関)
データの寄託元 寄託手続き	カナダ統計局自身	カナダ統計局自身
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.statcan.ca/start.html http://www.statcan.ca/english/Talon/talonflashtext.htm	http://www.statcan.ca/english/Dli/governance.htm
アーカイブスタッフ	不明	http://www.statcan.ca/english/Dli/statcan.htm

諸外国の統計データの二次的利用の状況

国名	カナダ	
データアーカイブ機関等 [主な提供データの名称]	カナダ統計局 (Statistics Canada)	
	RDC (Research Data Center)	Remote Data Access
機関又はデータの概要	<p>○カナダ国内に14箇所のRDCを設置 連邦政府の分析要望に応えるために、Federal Data Access Centreがオタワ大学に設立された。他の大学にはオタワ大学への専用の接続端末がある。 ○位置付け The university-based centres are, essentially, extensions of Statistics Canada offices (大学に設置されたセンターは、実質的に統計局の一部とみなされる)</p>	カナダ統計局のリモート集計方式で、許可された研究者が集計用のプログラムをインターネットを通じてカナダ統計局に送付し、結果を送り返してもらおうオンラインサービス。
運営経費の負担 (スポンサー)	カナダ統計局と大学間コンソーシアムとSocial Sciences and Humanities Research Council (SSHRC)により成る機関	カナダ統計局
マイクロデータ提供開始時期 又はアーカイブ設立時期	1998年	2002年4月に試験的プロジェクトとして発足
提供されるマイクロデータ	<p>Canadian Community Health Survey (CCHS) Ethnic Diversity Survey (EDS) Longitudinal Survey of Immigrants to Canada (LSIC) National Longitudinal Survey of Children and Youth (NLSCY) National Population Health Survey (NPHS) Program for International Student Assessment (PISA) Survey of Labour and Income Dynamics (SLID) Workplace and Employee Survey (WES) Youth in Transition Survey (YITS) Other Surveys</p> <p>主に完全な個票を提供</p>	<p>1998年 Adult Education and Training Survey 2000年 Program for International Assessment 2000年 Youth in Transition Survey</p> <p>(現在の状況)</p>
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	<p>Data Access and Control Service Division (アクセス・コントロール管理部門) 秘匿処理に関しては、前項参照</p>	同左
秘匿方法	<p>オンラインでは原則として秘匿を施していないマイクロデータ (契約時は連邦職員として扱われ、宣誓の義務を負う) PUMFの秘匿方法については前頁を参照</p>	秘匿されているが、利用者に結果を返送する際に、結果に特定因子が無いかを統計局で確認(見つければ結果表に秘匿を施す)
利用条件・手続	<p>SSHRCとカナダ統計局管轄の委員会に必要事項を記述した研究計画を提出 研究開始前にSSHRCへアクセスID(内部用番号)とパスワードを受領 Direct Access プロジェクト申請が承認されれば、連邦政府の職員として扱われ、データアクセスが可能となり、結果も統計局に提出する義務を負う(1か月のオンライン接続料は3500ドル) Indirect Access カナダ統計局と利害関係のある連邦政府の'Memorandum of Understanding'(覚書)が必要となる。特定の研究機関の合成ファイルが格納されているワークステーションにアクセスをする。結果も研究機関へ提出する義務を負う(1か月の間接的オンライン接続料は5500ドル) ○完全個票研究契約書(MICRODATA RESEARCH CONTRACT) オンライン利用限定</p>	<p>電子媒体にて研究計画書を提出 承認を受けた後、利用者は類似データを使用してプログラムテストを行う(SAS, SPSS, STATA and WESVAR) テスト完了後、オンラインにてプログラムをカナダ統計局へ送付して、結果を受け取る (必要に応じて秘匿処理を施されるが、利用者には通知される)</p>
提供形態	データセンターの端末で利用 (上記の利用法参照)	-
提供フォーマット	-	-
マイクロデータの利用料金	<p>(上記の利用法参照)</p> <p>Direct Access 1か月のオンライン接続料は3500ドル</p> <p>Indirect Access 1か月の間接的オンライン接続料は5500ドル</p> <p>PUMF:の提供も行っている</p>	利用後3~6ヶ月間が経過してから、少額の請求がある (金額は不明)
データの寄託元 寄託手続き	<p>カナダ統計局 (統計局よりオタワ大学に一括してデータを預けられる)</p>	<p>カナダ統計局 (統計局でプログラム受理、集計を行う)</p>
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	-	-
アーカイブスタッフ	カナダ国内のRDC関係大学の職員	カナダ統計局

諸外国の統計データの二次的利用の状況

国名	ドイツ	
データアーカイブ機関等 [主な提供データの名称]	ドイツ連邦統計局 (Federal Statistical Office Germany : Statistisches Bundesamt Deutschland)	
	PUFs (Public Use Files)	RDC (Research Data Centres)
機関又はデータの概要	PUFsはマイクロセンサス等から作成した匿名データ、 SUFsはより秘匿の緩い科学研究目的の汎用ファイルとして作成される。 (提供機関) 連邦統計局 州統計局	連邦統計局及び各州統計局に設置されているマイクロデータ利用研究所
運営経費の負担 (スポンサー)	内務省所管(連邦統計局) 首相直轄(州統計局)	同左
マイクロデータ提供開始時期 又はアーカイブ設立時期	1971年 マイクロセンサスからマイクロデータが学術目的に提供 1980年 連邦統計法改正(絶対的な匿名化措置) PUFsの提供開始 1987年 連邦統計法改正(事実上の匿名化措置) 匿名化の技術的・具体的検討 1997年 事実上の匿名化データ(Scientific Use Files)の提供開始	2001年に連邦統計局内に最初のRDCが設置される 次いで、2002年に各州統計局にRDC設置
提供されるマイクロデータ	・絶対的秘匿データ(Absolutely anonymised microdata) Public Use Files (PUFs) Campus Files (CF) ...for Universities and Schools (free) ・事実上の匿名化データ(Defact anonymised microdata) Scientific Use Files (SUFs) 上記のファイルを1利用者のために、特別に作成することもある	PUFs・CF・SUFsの提供のほかに、以下の方法によりデータ(或いは結果)を提供 ・Visiting Resercher Desktop(オンサイト利用) ・Controlled Remote Data Processing(リモート集計) ・Special Data Processing(オーダーメイド集計)
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	現在は不明 1988～91年にかけて「事実上の匿名化」の具体的検討が連邦・州統計局、データ保護 コミッショナー、マンハイム大学及びマンハイム社会調査方法論分析センター(ZU MA)にてなされた	同左
秘匿方法	Measures to anonymise microdata ・Perturbation(攪乱) ・Arbitrary arrangement of data records(データ記録の任意調整) ・Distribution of obsolete data(古いデータ配布) ・Omitting(削除) ・Presenting in less detail(詳細排除) (秘匿プロジェクトにZUMAも参加)	同左
利用条件・手続	PUFsとSUFsは個人のPCで利用可能(データは有料) Campus Files(CF)は学校等を対象とし、RDCにて無料提供している (CFはPUFsの一種だが、秘匿方法の違いは不明、学生が自由に利用できることを 考えると、PUFs以上に強い秘匿処理が施されていると思われる) 1987年の「連邦統計法」制定より、データへのアクセスが始まった	オンサイトについて、以前は「1ドル職員」(One-Dollar-Man)制度があり、1ドルが 任命費として臨時職員(公務員)化する その後、RDCで「デスクトップ客員研究員」制度が確立され、秘匿の緩い個票を利用 。期間中は施設内で固有パスワードを使い、データ保存も出来る(臨時職員であ るかは明確ではないが、宣誓を行う) リモートはSAS,SPSS、Stata用プログラムをRDCへ提出 オーダーメイドはRDCへデータ処理を全面依頼 いずれも結果表のみを受け取る
提供形態	RDCにて行われる (詳細は右記参照)	CD-ROM オンライン (CFはオンラインのみで提供)
提供フォーマット	-	-
マイクロデータの利用料金	RDCにて行われる (詳細は右記参照)	料金は不明だが、PUFsであってもライセンスと購入料金がかかる (CFのみ無料で提供され、ライセンスも必要ない)
データの寄託元 寄託手続き	連邦・州統計局	同左
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.destatis.de/cgi-bin/e_wwwwais	-
アーカイブスタッフ	-	-

諸外国の統計データの二次的利用の状況

国名	ドイツ	
	ZUMA	ZA
データアーカイブ機関等 [主な提供データの名称]	Centre for Survey Research and Methodology	Central Archive for Empirical Social Research, University Cologne
機関又はデータの概要	マンハイム調査分析センター	institute of the Cologne Association ケルン協会内の組織 1987年に計量歴史社会調査データのアーカイブを独立した機関として、設立された。 (設立以前はケルン大学の附属機関として活動)
運営経費の負担 (スポンサー)	1974年にマンハイムに設立。当初の予算の出典は不明。 1985年よりZA・IZと共に連合組織GESISとなり、連邦政府と11の州より予算を得ている。	1960年の発足時はケルン大学の組織として、大学より予算を得ていたが、現在は大学の組織から独立し、IZ・ZUMAと共に連合組織としてGESISより予算を得ている。 連合化以前より、ZUMAとZAのスタッフは主としてケルン大学出身であり、相互の関係は密接。
マイクロデータ提供開始時期 又はアーカイブ設立時期	1989年より連邦統計情報システム(STATIS-BUND)によりマイクロセンサスの提供を開始 そのほか所得・支出調査、家計パネル調査、行動調査等	1960年より、調査データ保管・整理・保管を始めた、欧州最古のアーカイブ機関。 肥大化を防ぐために、1969年に「情報センター(IZ)」を独立機関として設立した。
提供されるマイクロデータ	マイクロセンサスの1%抽出中の70~80%のリサンプリングデータ その他、科学調査や雇用統計も1%リサンプリングで提供されている 「事実上の秘匿」により、病院統計、建物・住宅センサス、道路交通事故統計なども提供されている	Face-to-Face Interviews Mail Surveys Telephone Surveys Online Surveys
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	不明	不明
秘匿方法	Measures to anonymise microdata ・Perturbation (攪乱) ・Arbitrary arrangement of data records (データ記録の任意調整) ・Distribution of obsolete data (古いデータ配布) ・Omitting (削除) ・Presenting in less detail (詳細排除) 過去の統計局の「秘匿プロジェクト」にZUMAも参加していたため、同様と考えられる	ZAは秘匿化業務を行わず、ZUMAの匿名標本を保管すると考えられる。
利用条件・手続	学術研究のため	同左
提供形態	オンライン CD-ROM	オンライン CD-ROM
提供フォーマット	-	-
マイクロデータの利用料金	科学目的で1サンプル130マルクで提供	German General Social Survey (ALLBUS) を実施し、社会科学の学界へ無料で調査データを公開している 基本的にZUMAメンバーは無料
データの寄託元 寄託手続き	連邦統計局	ZUMAの実施した調査 民間の調査機関の調査 (一部統計局の調査も含む)
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	-	-
アーカイブスタッフ	http://www.gesis.org/en/staff/zuma/zuma_mitarbeiter.htm	http://www.gesis.org/en/staff/za/za_staff.htm http://www.gesis.org/en/staff/iz/iz_mitarbeiter.shtm

諸外国の統計データの二次的利用の状況

国名	イギリス	
データアーカイブ機関等 [主な提供データの名称]	UKDA The U K Data Archive	CCSR (Cathie Marsh Centre for Census and Survey Research) [SARs (Samples of Anonymised Records)]
機関又はデータの概要	エセックス(ESSEX)大学内に設置 2003年1月よりESRCとJISCの共同出資で国家的なデータアーカイブ・提供センターであるthe Economic and Social Data Service (ESDS)がサービスを開始した。なお、ESDSは主要4アーカイブ機関と協働してデータの保存・提供のサービスを実施するが、ESDSのあらゆる方針と運営に関する責任はUKDAが負っている。	マンチェスター(Manchester)大学内に設置 1992年に ESRCの助成を受けて、Census Microdata Unit(CMU), the Centre for Census and Survey Research(CCSR)として発足、1995年に創立者のCathie Marsh教授の名を冠し、発展的に改組。 ESDSにおいては「ESDS Government」の主導的な役割(UKDAは支援)を果たす。
運営経費の負担 (スポンサー)	現在はthe Economic and Social Research Council (ESRC)とthe Joint Information Systems Committee (JISC) of the Higher Education Funding Councils が共同出資。(注)英国の学術研究機関への予算の提供は、教育雇用省ルートのHEFC(Higher Education Funding Council)と科学技術庁系統のESRCがあり、大学の教育関係予算は前者から、研究予算は後者から提供され、かつてはArchiveの運営も後者から提供されたが、1999年に HEFCの下にJISCが設置されてから、両者が半額を負担している。	the Economic and Social Research Council (ESRC)の助成を受けて発足(現在も助成を受けているどうかは不明)
マイクロデータ提供開始時期 又はアーカイブ設立時期	1967年	1992年4月
提供されるマイクロデータ	・社会科学・人文科学に関する5000以上のサーベイデータを提供。 ・1980年頃から政府統計データを収集(標本調査が中心、ただし、2001年センサスの世帯SARであるSL-HSARも)は「ESDS Government」の中で提供(CCSRが主導、UKDA支援)。なお、アーカイブ機関側でSARsデータのように変数の追加は行っていない。 (注)SL-HSAR(Special Licence Household Sample of Anonymised Records)は完全な秘匿ではないが詳細な情報が得られる特別仕様の世帯SAR(2001年のみ)。	1991年と2001年のセンサスから作成された匿名標本データ(Samples of Anonymised Record)を提供、個人と世帯の2種類SARがあり、年次により仕様が異なる。(ただし、2001年世帯SARはUKDAで提供)。このほか、5%抽出の2001年SAM (the 2001 Small Area Microdata file)は2006年より提供を開始。なお、より詳細な世帯・個人データのCAMS (Controlled Access Microdata Samples)はONSの施設内で利用可。 ・2%(3%)個人SARは約110(175)万レコードで、センサスにおける全個人項目と一部の世帯項目を含み、地域区分は278区分(北アイルランドを加えUK全体で288)。 ・1%世帯SAR('01はSL-HSAR)は約20万レコード(世帯員数約50万人)で、全世帯項目とCMUが追加した世帯・家族の変数を含み、地域区分は12区分('01は区分なし)。 ・5%SAMは約300万人の個人レコードで、地域は地方自治体レベルと詳細であるが、個人SARに比べ個人項目の情報は少ない。
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	隔週ごとに開催される収集検討委員会(ARC: The Acquisitions Review Committee)が具体的な基準に照らして寄託データ提案書の審査を行う。	CAPRI (Confidentiality and Privacy) ・データ収集、提供、分析における秘密保護に関する研究を行うチーム。現在のCAPRIは、Dr. Kingsley Purdamが主導し、マンチェスター大学中に構成員がいる。マイクロデータや小地域集計データの開示リスクのレベルを立証する革新的な方法を開発し、ONSのセンサスや標本調査の開示リスクの評価に多大なる貢献を果たすとともに、2005年5月の国際シンポジウムも成功裏に開催。
秘匿方法	SL(Special Licence)形式(完全な秘匿ではないが詳細な情報が得られる特別仕様のデータの提供は、現在ONSによって進められており、SL-HSARのほか、Annual Population Survey、Labour Force Surveyでも実施している。	センサスマイクロデータ提供に先立ち、ESRCに設置された故Cathie Marsh教授を委員長とするWGで、「SARによって秘密が露見する危険性は無視できる。」(91)と結論。 [5つの秘密保護措置] 低い抽出率(世帯1%、個人2%) データの圧縮(suppression)・・・希少・固有事例の除去(統合) top coding又は分類の統合(年齢、就業時間、産業、職業、教育など) 地域情報の制限(個人は広域地方自治体を、世帯はRegister Generalがベース) 露見リスク評価研究(Euからの資金援助を受けてモデルの開発に従事)
利用条件・手続	イギリス国内における学術研究機関の研究生・学生、政府機関の従事者は、インターネット上のAthens認証システムにログインすることによって、UKDAに保管されているサーベイデータを利用することが可能。 商業目的の利用者であっても、AthensのUser IDを保有していれば、データを取得可能 なお、2001年SL-HSAR等の利用にはUKDA経由でONSからSpecial Licenceの取得が必要。 政府機関もUKDAを通じて他の政府機関のデータをしばしば利用している。	1991SARs、2001SARs及び2001SAM (Licenced) (1)イギリス国内において非営利かつ学術目的で利用する研究者・学生は、AthensのUser IDとパスワードを取得した上で、Census Registration Service(CRS)に登録。 (2)非学術目的で利用する者は、End User Licence Agreementに記入し、署名した上で、契約書を郵送する。 個人CAMS Census Research Access Board(CRAB)の審査を受けた後に、CRABの推薦を受けて、ONSのマイクロデータ公開委員会(Microdata Release Panel)、The General Register Office for Scotland(GROS)及びThe Northern Ireland Statistics and Research Agency(NISRA)によって審議されたのちに、利用可能。 (注)国外の者は、従来、英国の研究機関又は高等教育機関の研究者となる以外は利用不可であったが、ONSは3%個人SARと5%SAMデータに限り、利用を承認。
提供形態	UKDAのWeb上でダウンロード Nesstar Webview FTPによる転送 CD-ROMの購入	1991年SARs(個人、世帯)および2001年個人SARs (1)Web上でダウンロード (2)Nesstarの利用 (3)NSDstatの利用 (4)CD-ROMの購入(30ポンド) 個人CAMS 利用申請が許可された場合に、ONS内で利用可能
提供フォーマット	SPSS Stata SAS タブ区切り(tab delimited)のファイル形式	SPSS Stata SAS カンマ区切りファイル 等
マイクロデータの利用料金	非営利目的でデータを取得する利用者については、無料でダウンロード可能 ・営利目的でデータを利用する場合は、データセット1つにつき500ポンドの手数料が発生 CD-ROMにかんする料金については、メディア代が7.50ポンド、handling feeが2.50ポンド、郵送料がイギリス国内で3ポンド(イギリス以外のヨーロッパ諸国では4ポンド、それ以外では5ポンド) 税率17.5%の付加価値税(Value Added Tax)が課せられるが、(1)教育目的で使用する研究生と学生、(2)VAT numberを付与されているEU諸国の利用者、(3)EU外部の利用者、部局内の目的(departmental purpose)で利用するエセックス大学の大学関係者は、税金を免除されている。	1991年SARs(個人、世帯)および2001年個人SARs(Licenced) (1)学術目的で利用する教育機関の研究者、学生は、センサスデータを無料で取得することが可能(CD-ROMを購入する場合には、30ポンドの経費がかかる)。 (2)商業目的の利用者については、データセット1つにつき1000ポンドの経費を要する。 (3)地方政府、保健機関、中央政府と慈善団体の従事者については、同500ポンド。 1ライセンスで10人まで利用可能(追加料金で増員可) (2)と(3)については、さらに付加価値税(VAT)が課税される。 個人CAMSについては、1日あたり52ポンド+VATの経費が生じる。
データの寄託元 寄託手続き	大学等の民間の研究機関(者)、政府機関 UKDAのacquisitions teamは、ESDSの援助の下で、調査のあらゆる段階でデータ作成者や寄託者の支援に専念している。(データ作成者への支援と寄託手続は下記のとおり) Advice for data creators - introduction Depositing data - introduction	現在は国家統計局(ONS: the Office for National Statistics) 以前は人口センサス調査局(the Office of Population Censuses and Surveys)
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.data-archive.ac.uk/ http://www.data-archive.ac.uk/about/about.asp Data Management	http://www.ccsr.ac.uk/
アーカイブスタッフ	UKDA Staff (UKDAとESDSのDirectorは同一人物)	Staff Member

諸外国の統計データの二次的利用の状況

国名	イギリス	
データアーカイブ機関等 [主な提供データの名称]	CeLSIUS (Centre for Longitudinal Study Information and User Support) [LSデータ(Longitudinal Study Data)]	ONS (the Office for National Statistics) IDBR (Inter-Departmental Business Register)
機関又はデータの概要	ロンドン(LONDON)大学内に設置 CeLSIUSはロンドン大学のLSHTM (the London School of Hygiene and Tropical Medicine)の人口研究センター(Centre for Population Studies)内に設置された。	ビジネスレジスター部門 (Business Register Unit) IDBRは関係部局間で使用できるビジネス・レジスター
運営経費の負担 (スポンサー)	ESRCとJISCの人口センサプログラムによる共同出資(同プログラムは2006年7月31日まで)。	
マイクロデータ提供開始時期 又はアーカイブ設立時期	2001年12月(CeLSIUSの発足) ・LSデータの提供は1982年よりロンドンシティ大学社会統計研究部門(SSRU: the Social Statistics Research Unit)のLS Support Programmeを通じて、その後1998年10月よりロンドン大学のCentre for Longitudinal Studies、さらにCeLSIUSに引き継がれ	
提供されるマイクロデータ	・LSデータ(Longitudinal Study Data)は、リンケージされた1%抽出(約50万)のセンサス個票データ(England,Wales)である。1974年当初は、職業別死亡データの妥当性や長期的な出生データの欠如という問題に対処するために作成されたが、以降、様々な研究課題を処理するために用いられてきた。(スコットランドのLSは現在作成中) 特定の4つの誕生(月)日に生まれた人をLSサンプルとして、1971年のセンサス、以降のセンサスや他の登録情報とのリンケージによってサンプルの追加や既存サンプルへのイベント情報(死亡、サンプルとなっている母の出産、移住、癌患者登録など)の追加が行われる。	IDBR (Inter-Departmental Business Register)は統計目的のビジネスレジスターである。ONSや他の官庁に企業を対象とした標本調査のサンプルフレームの提供するとともに、企業活動の分析の主たるデータ源となっている。 ・IDBRを用いた標準的な分析 - Disclosive data - Local Authorities (disclosive data) - Non-disclosive data - Samples (disclosive data)
マイクロデータ提供に関する審議機関 (開示方針、秘匿方法等の検討・決定機関)	LSRB (the ONS LS Research Board) Criteria for the approval of applications to use the LS data	Micro Data Release Panel (MRP) データの秘匿方針はNational Statistics Code of Practice (Protocol on Data Access and Confidentiality)に基づく。
秘匿方法	個人レベルのデータは決して提供されない。 公表の際は、公開リスクを減じ、研究とデータが適切に説明され、かつ、利用されていることを保証するため'final outputs clearance'という手続がある。主な基準は、 ・表章結果は1セル当たり最低3標本の合算値とすること。 ・LSについて正しく表現されていること、標準的な記載例は利用可。 ・表や数値には適当な場所に、「原典: ONS Longitudinal Study」と記記。 ・LSに含まれるデータはイングランドとウェールズのみであると結果を正しく評価。	
利用条件・手続	[学術機関の利用者の場合](学術機関以外はONSのLS課に直接コンタクト) 自分の研究の論点を明確にしておく LSデータが自分の研究にマッチするかHPで確認 CeLSIUSの職員にコンタクトを取り、研究課題、研究期間について相談 オンライン・トレーニングを実施 秘匿(Confidentiality)に関する契約書(ONS LS Licence Agreement)を読む 研究計画書(Application form A)のe-mailで送信(又は郵送) CeLSIUS職員の指示に従い、Data Dictionaryを用いて、データや変数の抽出・加工の申請書(Application form B)を作成、ONS LS Research Board(LSRB)に提出 LSRBによるLSデータ利用承認後、契約書(ONS LS Licence Agreement)の提出 当該研究を担当するCeLSIUS職員から合算データか、分析の実行結果を受領(中間結果は推奨暗号化ソフトのSafeGuard PrivateCryptを用いて暗号化して提供) 発表しようとする際は、事前にLSRBからfinal outputs clearanceを取得しなければならず、また、CeLSIUSとONSに対する謝辞を明記。	{IDBRを利用できる者} ・中央政府の職員 - 調査及び分析のため情報を利用可。 ・地方政府及び他の政府団体の職員 - 一部のdisclosive dataを利用可。 ・政府と契約した者 - 契約で特定の調査及び分析のため情報(を利用可)。 ・一般大衆 - disclosiveでないものであれば、分析のため情報を利用可。 [disclosive dataの利用] 公的に権限を与えられた外部の団体や政府機関で働いている契約者のみで、下記の確認書に署名が必要。 データはONSと合意した目的に対してのみ利用されること。 データの安全性が確保されていること。 ONSの許可なく第三者にデータを開示(disclose)しないこと。 違法開示に対しては法的制裁が科されること。
提供形態	オーダーメード集計(通常、ONS(又はCeLSIUS)が集計した統計表や摘要表(合算したデータを含む)などが利用申請者に渡される。ただし、1、2の簡易な表であれば最小の公的手続のみでCeLSIUSが対応することもある。なお、Multi-level Modellingの実行など何らかの事情があれば、ONSに直接出向いて同データを用いた分析も可能... オンサイト利用)。 LSデータの公式の利用窓口として、下記の二つがあり、両者は緊密に連携し、利用者のサポートを行っている。 ・CeLSIUS.....大学等の学術機関向け ・ONSのLS Unit (LS課).....上記以外の機関向け	標準的な分析であれば、下記の期間内にオーダーメード集計結果(又は匿名化されたデータ)として提供される。 - Disclosive data...Micro Data Release Panel (MRP)の承認の確認から、2週間以内 - Local Authorities (disclosive data)...作業の確認をしてから、2週間以内 - Non-disclosive data...作業の確認をしてから、2週間以内 - Samples (disclosive data)...作業の確認をしてから、4週間以内
提供フォーマット	(原データは提供されず、結果表や合算データが提供される)	ハードコピー ディスク CD (Disclosive dataはe-mailとfaxでの送付は不可)
マイクロデータの利用料金	イギリス国内の学術機関における教職員・学生に関しては、無料で提供	・オーダーメード集計の利用料金(ONSが請求) - Disclosive data £ 60/hour(必要に応じ付加価値税も加算) - Local Authorities (disclosive data) £ 150/district(同上) - Non-disclosive data £ 60/hour(同上) - Samples (disclosive data) £ 60/hour(同上)
データの寄託元 寄託手続き	(データはONSがDBで保管している。CeLSIUSのスタッフは唯一DBにアクセスし、データ抽出を行える。)	IDBRの主たるデータベースは、430万企業中の210万企業をカバー。 付加価値税(VAT: Value Added Tax)システムの商業者...HMRC 1より毎日提供 源泉徴収(PAYE: Pay As You Earn)システムの雇い主...HMRCより四半期毎に提供 Companies House 2に登録された法人企業 このほか、ONS Business Register Survey等の調査でこれらの行政記録を補う
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.celsius.lshtm.ac.uk/	http://www.statistics.gov.uk/ http://www.statistics.gov.uk/CCI/nugget.asp?ID=981&Pos=&ColRank=1&Rank=374
アーカイブスタッフ	http://www.celsius.lshtm.ac.uk/staff.html	連絡先 mailto:carl.roberts@ons.gov.uk mailto:IDBRDAS@ons.gov.uk

諸外国の統計データの二次的利用の状況

国名	オーストラリア	
	オーストラリア統計局 (Australian Bureau of Statistics) [CURFs (Confidentialised Unit Record Files)]	ASSDA Australian Social Science Data Archive
機関又はデータの概要	CURFsはABSの主要統計調査の匿名標本データの総称	
運営経費の負担 (スポンサー)		
マイクロデータ提供開始時期 又はアーカイブ設立時期	1984年(国民健康調査のデータ)が最初 本格的な提供は1990年代から	1976年設立(RSSSの付属機関) 2001年よりACSRに吸収される
提供されるマイクロデータ	<ul style="list-style-type: none"> ABSで調査される諸調査(表1参照)で、センサスをはじめとする60調査程度(ホームページ掲載数) 基本的なデータ(Basic)とより詳細なデータ(Expanded)が選択可 	<ul style="list-style-type: none"> 大学や民間機関等の社会調査や世論調査 近隣諸国の諸調査 など1500程度
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)		
秘匿方法		
利用条件・手続	<ul style="list-style-type: none"> ABSによる2週間以内の審査を通過した者が利用可能(原則として国内からのみ申請可能、例外あり)。 申請には、Contact OfficerとResponsible Officerを定め、統計目的以外に利用しないなどと誓約 	<ul style="list-style-type: none"> 申請書を提出した後、誓約書を提出する。ほとんど誰でも利用可能 調査データの寄託者から利用者に対する付加的な制約があるデータもあり
提供形態	<ul style="list-style-type: none"> CD-ROM RADL(Remote Access Data Laboratory)・・・インターネット経由で直接分析、各大学で利用可能 ABSDL(ABS Site Data Laboratory)・・・ABS内のみで利用可能 	<ul style="list-style-type: none"> CD-ROM インターネット経由でダウンロード NESSTAR Light(ウェブ上で統計分析も行う)
提供フォーマット	SAS SPSS Stata	SPSS ASCII SAS
マイクロデータの利用料金	<ul style="list-style-type: none"> 1996年1月1日以降のデータ 8,000豪ドル 1986年1月1日～1996年6月30日のデータ 5,500豪ドル 1986年1月1日以前 1,100豪ドル(いずれもGSTを含む) ABS-AVCC CURF 協定に加盟している大学からの研究目的の利用は無料 	<ul style="list-style-type: none"> 基本料金60豪ドル+処理費用+ディスクなどの実費+郵送料+GSTで、だいたい100豪ドル以内 ACSPRIの加盟機関の場合は割引あり
データの寄託元 寄託手続き	基本的にABSの実施する調査	ASSDA外の大学研究者や調査機関からの寄託を募る
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.abs.gov.au/websitedbs/d3310114.nsf/Home/Home?OpenDocument	http://www.assda.asn.au/asp/index.asp
アーカイブスタッフ		

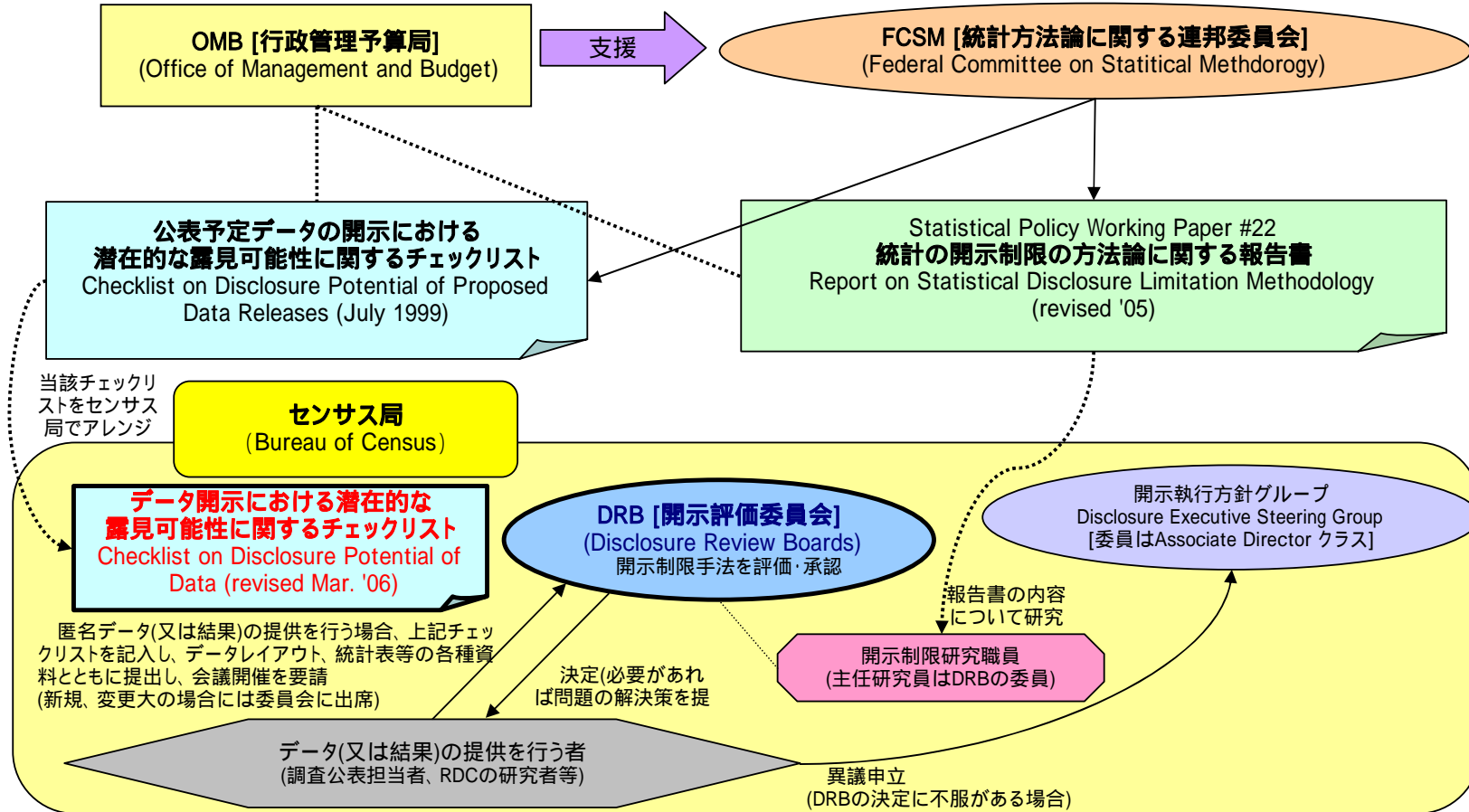
諸外国の統計データの二次的利用の状況

国名	ニュージーランド
データアーカイブ機関等 [主な提供データの名称]	NZSRDA NEW ZEALAND SOCIAL RESEARCH DATA ARCHIVES
機関又はデータの概要	マッシー大学社会科学学部
運営経費の負担 (スポンサー)	
マイクロデータ提供開始時期 又はアーカイブ設立時期	・1992年(マッシー大学の社会科学学部)に設立
提供されるマイクロデータ	・国内の社会科学データが33種程度 ・海外のデータが2種
マイクロデータ提供 に関する審議機関 (開示方針、秘匿方法 等の検討・決定機関)	
秘匿方法	
利用条件・手続	・データリスト(国内、海外)の中から、データセットを選択する。 ・申請書を提出する。
提供形態	・電子メールでの添付ファイル ・普通郵便による郵送
提供フォーマット	不詳・・・利用者からの要望に応じる?(明確な記述が無い)
マイクロデータの利用料金	・基本料金は媒体料金 ・利用要請が政府、商業目的、教育関係により段階的な設定 ・データセットのサイズによる加算料金を設定 ・詳細は、利用申し込み後に渡される。
データの寄託元 寄託手続き	
当該サイトのトップページ Mission Statement 年次報告書(事業報告書)	http://www.massey.ac.nz/~nzsrda/nzsrda/archive.htm
アーカイブスタッフ	

【参考引用文献等】
 「2004年度 統計データアーカイブに関する調査 研究報告書」(平成17年3月
 (財)統計研究会 総務省統計局委託研究)の「各国データアーカイブについての比較表」
 「講座マイクロ統計分析 統計調査制度とマイクロ統計の開示」(松田芳郎、濱砂敬郎、
 森博美編著 日本評論社 2000年)
 各国統計局、統計データアーカイブ機関のホームページ(2006年9月時点)

(注)本資料は を基に、及び から得た情報を付加して作成したものであるため、
 各資料における把握時点、情報量の違いにより、掲載内容が統一されず、必ずしも
 整合的でない可能性があることをご了承願いたい。

米国における秘匿処理の審査について



なお、下記の機関においても、DRBを設置しており、承認基準はセンサス局と同様。

- 国立保健統計センター
(National Center for Health Statistics)
- 労働統計局
(Bureau of Labor Statistics)
- 国立教育統計センター
(National Center for Education Statistics)

資料「Panel on Disclosure Review Boards of Federal Agencies (Presented at the Joint Statistical Meetings August 17, 2000 Indianapolis, Indiana)」に基づいて、独立行政法人統計センターで整理

注) 組織名・文書名の訳語は本資料のために便宜付けたものである。

Disclosure Review Boards の構成

Bureau of Census (センサス局)			
目的・職責	<p>[DRB綱領] 公衆に利用されるすべてのデータ、成果物に関して、センサス局の開示制限方針を確立し、評価すること 公衆に利用されるすべてのデータ、成果物の公表のための開示制限手続の評価・承認をすること 前記の方針の枠組みの中で、DRBは開示制限方針をプログラム責任者、センサス局職員、データ利用者と将来のスポンサーに対して伝えること ミクロデータと結果表における開示の可能性や開示制限技術の有効性に関しての教示と調整 必要に応じて、センサス局の開示方針の改定</p>		
委員の人数	9名の投票選出委員と数名の交代制委員(任期3年)		
	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; border: none;"> 常任委員3名(うち1名は委員長) [常任委員は投票選出委員] ・委員長: 統計研究部の上席研究員 ・事務局: 方針策定事務局(Policy Office)代表 ・事務局: 統計研究部の主任研究員 </td> <td style="width: 50%; border: none;"> 非常任委員6名 ・人口統計部(2名) ・センサス部(1名) ・経済統計部(3名) (うち1名はCESとRDCの監視) </td> </tr> </table>	常任委員3名(うち1名は委員長) [常任委員は投票選出委員] ・委員長: 統計研究部の上席研究員 ・事務局: 方針策定事務局(Policy Office)代表 ・事務局: 統計研究部の主任研究員	非常任委員6名 ・人口統計部(2名) ・センサス部(1名) ・経済統計部(3名) (うち1名はCESとRDCの監視)
常任委員3名(うち1名は委員長) [常任委員は投票選出委員] ・委員長: 統計研究部の上席研究員 ・事務局: 方針策定事務局(Policy Office)代表 ・事務局: 統計研究部の主任研究員	非常任委員6名 ・人口統計部(2名) ・センサス部(1名) ・経済統計部(3名) (うち1名はCESとRDCの監視)		
会議の開催	原則として毎週 (ただし、定足数に満たない場合又は議題が無い場合は開催しない。)		
評価の対象	<p style="text-align: center;">結果表、PUMS(Public Use Microdata Samples)、プログラムの全般を評価</p> <ul style="list-style-type: none"> ・人口センサスの結果表の仕様 ・経済統計のセル秘匿プログラムの仕様及びパラメータ ・最新と旧式のPUMSの仕様とデータレイアウトとの比較評価 ・DRB開示制限ガイドラインを満たした特別結果表 		
審議手順等	<p style="text-align: center;">チェックリスト(Checklist on Disclosure Potencial of Data)等に基づく審査</p> <p>調査実施部から送付されたセンサス局開示チェックリスト、調査票、重要な変数の一覧表、ミクロデータのデータレイアウト、結果表の概要、標本設計識別因子(可能であれば)を受領し、その他課題に関連しそうな情報についても審議を行なう。委員長は新規の申請又は大きな変更の場合など、必要に応じて調査担当者を会議に招集し、行なった処置の妥当性等の説明と質疑応答の機会を与える。その後、委員間で露見リスクの評価について審議が開始される。</p> <p>その他、近年、秘匿措置の審議に係る手順の効率化を図るため、データ収集及び集計処理の前に開示問題の解決を図る試みがなされている。</p> <p>委員会の最終決定は多数決により決められる。申請が却下された場合、センサス局内の“開示執行方針グループ”(Disclosure Executive Steering Group)へ異議申立できる仕組みとなっている。</p> <p><u>チェックリストについては、データ提供につきものの開示と秘匿の問題について、いかなる組織にも教育し、認識させることができる非常に有益な道具として評価している。</u></p> <p style="text-align: center;">開示執行方針グループ(Disclosure Executive Steering Group)の構成員</p> <ul style="list-style-type: none"> ・ Principal Associate Director for Programs ・ Associate Director for Methodology and Standards ・ Associate Director for Demographic Surveys ・ Associate Director for Economic Censuses and Surveys ・ Associate Director for Decennial Censuses 		
研究業務	<p style="text-align: center;">開示制限研究職員(Disclosure Limitation Research Staff)</p> <p style="text-align: center;">周期センサス、人口統計、経済統計の各々における課題への対応を主とし、Statistical Working Paper #22 (OMB) について以下を対象とした内部研究を実施</p> <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; border: none;"> 主任研究員(DRB委員) ・ 地理情報の閾値の決定 ・ トップ及びボトムコーディング ・ ランダムスワッピング ・ 再格付 ・ その他、外部研究者により実行されている方法の評価・研究 </td> <td style="width: 50%; border: none;"> 他の研究員 ・ 分布評価 ・ 区分統合 ・ セル秘匿 </td> </tr> </table>	主任研究員(DRB委員) ・ 地理情報の閾値の決定 ・ トップ及びボトムコーディング ・ ランダムスワッピング ・ 再格付 ・ その他、外部研究者により実行されている方法の評価・研究	他の研究員 ・ 分布評価 ・ 区分統合 ・ セル秘匿
主任研究員(DRB委員) ・ 地理情報の閾値の決定 ・ トップ及びボトムコーディング ・ ランダムスワッピング ・ 再格付 ・ その他、外部研究者により実行されている方法の評価・研究	他の研究員 ・ 分布評価 ・ 区分統合 ・ セル秘匿		
発足の歴史	<p>1981年 ミクロデータ評価検討会(Microdata Review Panel) 経済・人口の領域に分かれ、各々について独立評価を行っていた。</p> <p>1995年 開示評価委員会(Disclosure Review Board) 露見リスク回避を保証し、センサス局内で一定の開示制限方法を維持するために発足。</p>		

出典: Panel on Disclosure Review Boards of Federal Agencies(Joint Statistical Meetings)

注) 組織名・文書名の訳語は本資料のために便宜付けたものである。

参考資料(Supporting Document)

データ開示における潜在的な露見可能性に関するチェックリスト (仮訳)

バージョン1.2
2006年3月9日出版

センサス局標準
開示レビュー

Laura Zayatz 著
開示評価委員会 委員長
統計調査部

ドキュメント管理とコントロール

バージョン	出版日	承認	説明
1.0	2004年3月12日	Associate Directors	初版
1.1	2005年6月14日	Configuration Mgr.	Census Bureau Identity Standard and Quality Program Document Management Planに適合させた
1.2	2006年3月9日	Configuration Mgr.	主要標準のハイパーリンクを挿入

このドキュメントの最新バージョンは、センサス局のイントラネットで保守されており、品質管理保管所(the Quality Management Repository)を通じてアクセス可能です。

一旦このフォームに記入されたものは、センサス局の機密事項になります

データ開示における潜在的な露見可能性に関するチェックリスト

あなたは、このフォームに記入する必要がありますか？

あなたの報告義務を削減するために、必ずしも繰り返し実施されるすべての標本調査又はセンサスの公表の度にこのチェックリストを完成させる必要はありません。もし、下記の基準にすべて合致すれば、あなたは、単にDRBの議長あてのメモを用意するだけでよい:地理的情報に変更がない、新規調査事項がまったく導入されていない、最初のデータ提供によって是認された開示回避手段がその後の全ての提供において履行されている、以前の提供においてチェックリストにすべて記入されていた。

たとえチェックリストがこの場合不要だとしても、データが再識別化(reidentification)に使用されかねない情報の公共利用に何らかの変化があることを考慮し、データは引き続き再検討が必要であることに注意して下さい。

標本地域が100,000人未満で地理的区域が識別されない人口統計調査に基づく表形式のデータを除き、どのような形式のデータ提供(例えば、人口統計又は経済統計のマイクロデータ、人口統計又は経済統計の表形式のデータ、オーディオ・テープなど)であっても、このフォームに記入しなければなりません。

このフォームはセンサス局の機密事項であることに注意して下さい。

データ開示における潜在的な露見可能性に関するチェックリスト

センサス/標本調査名: _____ 日付 _____
プロジェクト・マネジャー名: _____ 課. _____ 係. _____ Ph. _____
スポンサーの官庁: _____
提供予定時のデータの経年: _____ (年)

下記のうち、当てはまるものがあればチェックして下さい:

- [] 今回の申請は、一回限りのデータ作成のためのものである。
- [] 今回の申請は、本質的に同じ内容を持つ一連の提供のためのものである。
(次の提供がなされるまでの期間を具体的に記してください。)
- [] 今回の申請は、補足的あるいは以前発表していないデータについてのもので、承認を受けた成果物の再公開のためのものである。
(マークした場合、DRB/MRPに元の成果物を提出した日を記入して下さい)
- (答えが現在と異なるチェックリストの質問があればその質問にのみ、記入する必要があります。)

このチェックリストは3つのセクションに分かれています。該当するセクションについて、すべての質問にお答え下さい。もし、答えるのにスペースが足りない場合は、どうぞ紙を継ぎ足して、質問の番号が分るようにしておいて下さい。

- セクション1(2-10頁)は、マイクロデータに関する質問です。マイクロデータファイルは、回答者レベルのレコードで構成されています。各レコードには、人、世帯、事業所及びその他の構成単位にとって貴重な変数が含まれています。ほとんどのマイクロデータファイルは、人口統計的な情報を含みます。このセクションのいくつかの質問は、事業所ベースのファイルには適用できないかもしれません。
- セクション2(11-12頁)は、表形式(頻度計算)の人口統計データに関するものです。頻度計算データは、1セル内の構成単位(人、家族、など)の数を表します。このチェックリストは、センサスから作表されたもの、あるいは識別される地理的区域が標本数100,000人以下であれば、記入する必要があります。
- セクション3(13-16頁)は、事業所における表形式(規模)のデータに関するものです。規模のデータは、1セル内のあらゆる構成単位の利益の量の合計を表します。

セクション1 ミクロデータ

1.1. ファイルにおける地理的情報

原則: 識別される地理的区域には、標本地域内に最小でも100,000人いなければならない。

1.1.1 ファイル上で表示されるのは地理的にどのレベルですか?

ファイル上にある明白な地理的識別子だけでなく、データ項目、レコード識別子、あるいはファイル構造が、推測されうる余計な地理情報をもたらしていることもあります。したがって、これらの処置は最小人口100,000人の基準に適合しない地理的区域を不注意により識別されてしまうことを避けるために取られなくてはならないものです。問題が起り得る地域は、以下で吟味します。それぞれの地域ごとに、申請したファイルが提供される前に今まで行ってきた、あるいは、これから行われようとする活動について、どうか指摘願います。

- 1.1.2 第1 標本単位(PSU)あるいは他の地理情報は、内部利用のために設計された管理番号として組み込まれているのが一般的である。

提供されるファイルにおいて、どのようにこの問題が避けられるか:

___ 管理番号を削除するか、地理情報を含めない。

___ 管理番号を暗号化する;具体的に述べて下さい。

___ その他;具体的に述べて下さい。

- 1.1.3 多くのデータベース内のレコードは規則的に連続しているものであり、そのため、最初の事例は小さな番号の付されたPSUあるいはアルファベットの順序で最初になる郡となる。このファイルの記録がそのような地理的な推測が避けるように、どのような順番にならべられているか手短かに説明して下さい。

- 1.1.4 居住についての特定の地理(情報)を含意するデータ項目は、当委員会のために用意した人口統計表上で表示されている明白な識別子以上のものを明らかにし兼ねません。例:スペインの姓が含まれているか否か(5つの南西の州でのみ暗号化されている)という項目はそのグループの州であるという明白な識別子がファイルに存在しない場合であっても明らか; 地下鉄のある領域から地下鉄のない領域までの移動を識別する移動コードは、地理的識別子の一部として地下鉄の有無が含まれていない場合であっても明らか;原子炉または空港のXマイル内の住宅については、識別できる地域がたった1つだけでも明らか;電話局番;地域コード;あるいは、緯度と経度も同様に働きます。

この理由により削除されることになるすべての項目をリストアップして下さい:_____

他の項目において地理的な意味を持つ可能性はあると思うが、削除すべきかどうか決めかねているものがあれば、すべてリストアップして下さい。

-
- 1.1.5 サンプル情報はまた何らかの地理的な指標を提供する可能性があります。例えば、あるウエイトを見ればself-representingとnonself-representingのPSUsが識別される、あるいは意図的に過剰にした標本抽出地域の種類が明らかになる可能性があります。また、「ダービン・タイプ(Durbin Type)」、「ヒットナンバー(Hit Number)」などのためのコードは、地理に関係するかもしれません。

秘密保護の理由やウエイトを識別しにくくするような副標本プランのため、分散が推定されることから削除されるものがあれば、それらを含めてすべてのサンプル情報をリストアップして下さい:

他のサンプル情報において地理的な意味を持つ可能性はあると思うが、削除すべきかどうか決めかねているものがあれば、すべてリストアップして下さい。

- 1.1.6 ファイル上のどのデータ項目も、入手可能な情報を基に、ある特定の種類の施設(例えば刑務所または老人ホーム)の住居であることが明らかになることがあります。そのようなデータ項目はありませんか。その施設が識別される地域内でたった1つだけのようなケース、あるいは、そのような目的で記録する仕組みが得られているようなケースです。

はい - 施設のタイプを明示して下さい

いいえ

- 1.2. 個人(情報)の開示に係る特段のリスクをもたらすファイルの内容
公開用マイクロデータの開示基準は、申請された内容のいずれもが個人(情報)の開示に係る特段のリスクをもたらすかどうかを決定するため、それぞれのファイルのレビューを必要とします。DRBは、個々の回答(調査対象)者の秘密を保護するために取りうるいくつかの開示回避手段を明らかにしてきました。その手段は下記に述べたとおりであり、また、申請されたファイルに関連した重要な情報は、レビューの際の委員会(審議)の一助となるよう求められます。
- 1.2.1 氏名、住所、そのほか社会保障(制度)、高齢者医療保険制度又は低所得者医療扶助制度の番号といったユニークな数字の識別子は、ファイルから除外されなければなりません。
- 1.2.2 高収入は、個人又は世帯の明らかな特徴であり、また情報の中でも繊細な取り扱いが必要な項目と考えられています。したがって、ファイルにあるそれぞれの収入額は、世帯か、個人か、あるいは家族であるか、また、合計の所得か、個々の構成員の所得が含まれているかにかかわらず、トップコーディングされるべきです。母集団(人/世帯)全体に適用される収入変数のトップコーディングは、少なくともすべての事例の0.5パーセントが含まれるようにすべきです。
Subpopulations(特定の人口)に適用される収入変数の場合、トップコーディングは、適当

な事例の3パーセント、又はすべての事例の0.5パーセントのどちらかを含まれるようにすべきであり、どちらであってもかなり高いトップコーディングです。この原則の例外は、ある条件の下で起こりえます;例えば、地理的な詳細情報がほとんどない場合です。このようなトップコーディング原則の例外(的許可)はファイルの提供を承認する最終意見提出前に委員会で十分な議論がなされます。

すべての収入のトップコーディングは、適切な原則に合致していますか:

___ はい

___ いいえ - 何パーセントのトップコーディングとなっているか、トップコーディング(対象)の総数を明らかにし、委員会での審議を手短かに要約して下さい。

- 1.2.3 収入だけではなく、そのほかある特徴によっては、ある人と他人との違いがはっきり見えるようなものがあります。例えば、非常に高齢、自身の資産、賃貸料、住宅ローン総額が高価値あるいは高価格なケースです。ファイル上に示された地理の詳細によっては、このような項目が間隔又は順序を表す変数である場合には、これらにトップコーディング(及び/又は統合(collapsing))を施すよう考慮すべきです。当委員会は、(ウエイトを考慮して)ファイル上に代表される母集団(個人/世帯)全体の少なくとも0.5パーセントがこのようなトップコーディングの範疇として含まれることを提案します。ケースとしては少ないのですが、ごく少数の人口のみに適用される変数のような場合、当委員会は、トップコーディングの範疇として、適当なSubpopulations(特定の人口?)の約3~5パーセントを含めるべきであると考慮する場合があります。承認されたトップコーディングの例は:

年齢 - 90歳(2000年人口・住宅センサスの総人口の約0.5パーセント)

現金払いの居所(Housing Unit)の賃貸料 - 1,700ドル(2000年人口・住宅センサスの居所の約0.5パーセント)

暖房用燃料代 - 2,100ドル(2000年人口・住宅センサスにおいて、暖房用燃料が使われ、かつ、賃貸料や共益費とは別に支払われ、現在人が住んでいる居所の約3.0パーセント)

トップコーディング(又は統合(collapsing))及びトップコーディングに相当するものをすべてリストアップして下さい:

トップコーディングの必要性に関して疑問がある項目があれば、そのすべての項目をリストアップして下さい:

- 1.2.4 トップコーディングするデータ項目として提供されるべきであるとして何らかの提案があった場合は、その情報について詳述して下さい(例えば、トップコーディングした変数の平均値又は中央値)。

-
- 1.2.5 そのほかにも、個人を目に見えるようにするかもしれない特徴があります。それは、地理に影響されるものの、非順序変数として表示されており、それゆえトップコーディングできないものです。例えば、外国語、あるいはインドの部族の言語を示す符号;エスキモー人、アレウト族、グアム人、あるいはサモア人といった詳細な人種の識別子;前住居の場所の符号などです。このような場合、詳細値の総計はより大きいカテゴリーへ統合させる(collapsed)必要があるかもしれません。

秘密保護の理由で統合(あるいは削除)される項目があればすべてリストアップして下さい:

詳細を統合(collapse)する必要性に関して疑問がある項目があれば、そのすべてをリストアップして下さい:

- 1.2.6 コンテキストチュアルな(contextual)変数(その地域で人または世帯が住んでいることを文脈上物語っている変数)

どのようなコンテキストチュアルな変数があるか、また、暗号化されるレベルを明らかにして下さい。

秘密性理由で統合(あるいは削除)されるコンテキストチュアルな変数があればすべてリストアップして下さい:

詳細を統合(collapse)する必要性に関して疑問があるコンテキストチュアルな項目があれば、そのすべてをリストアップして下さい:

- 1.3. 外部ファイルとのマッチング能力に関する開示リスク

外部のファイルは通常住所と氏名を含んでおり、したがって調査回答(対象)者を識別するために用いることができることから、当該ファイル上のマイクロデータと外部ファイルのデータが一致する潜在的な可能性を回避するための努力なされなければならない。もし、他の行政機関や組織が整備しているメーリング・リストや行政記録上で見つけられる極めて具体的な特徴がその調査内容に含まれていれば、このように一致する可能性があります。

例えば、特定の地理的識別子と結合して自動車の製造元、モデル、年式が含まれていれば、これらの項目は氏名と住所を含む自動車の登録リストとのマッチングが可能なので、受け入れ難いものとなります。このような項目は広いカテゴリーとして再コードしなければ、たぶんファイルに残すことはできないでしょう。いくつかの例を挙げると：製造業者の所有する特定の主要耐久財(例えば、飛行機)の買い手のリスト;ある州の有権者登録リスト;連邦、州、あるいは地方自治体の納税記録;刑事裁判制度記録;州の狩猟及び漁労免許記録;ある同業組合の会員名簿など。

調査のための標本フレームがセンサス局以外のソースから来ている場合、マッチングはまた極めて可能です。標本フレームを提供した行政機関が調査データと元の記録を一致させることは、とりわけ調査記録に当該機関に由来するデータが含まれていれば、可能です。例えば、事業(プログラム)給付の受領総額、事業(プログラム)の開始年月日など。

1.3.1 外部ファイル

1.3.1.1 この申請されたファイルにあるデータを含む行政記録、メーリング・リストあるいは他の外部ファイルがあることをご存知ですか。

はい - そのリストを明らかにして下さい _____
 いいえ

1.3.1.2 申請されたファイルに含まれる標本がセンサス局以外のソースによって提供されたりリストから抽出されたものですか。

はい - ソースを明らかにして、どのように、また、誰によって標本がリストから抽出されたか記述して下さい。:

 いいえ

1.3.2 マッチング

外部ファイルが存在する場合、調査データとこのファイルをマッチングさせる可能性を避けるために、いくつかの手段が取られる可能性があります。例えば、選択した項目を、削除又は再コードする、あるいは、「ノイズ(誤差)」(すなわち、少量の乱数による変化)をこれらの項目に導入する可能性があります。当委員会は、マッチングの可能性を回避させるためにどちらの手段を取られなければならないか、あらかじめ、正確に示すことはできません。しかし、外部データベースとのマッチングの可能性が存在する場合、当委員会では、ファイル提供に関連したリスクを決定する際に、(次のような)いくつかの要因を考慮します: 1) マッチングする目的で利用できる変数の数; 2) マッチングを実行するのに必要な資源; 3) データの年齢(経年); 4) 外部データの入手しやすさ、信頼性、完全性; そして、5) データの内容が繊細に扱われるべきものか、あるいは独特なものか。マッチングをより容易にするいくつかの要因については下記にリストされています。また、マッチングの可能性を避けるためファイルが提供される前に取るべき手段として、情報が必要になります。(注: たとえあなたがマッチングに用いられるかもしれない外部ファイルとしてどのようなものがあるか知らないとしても、この情報は必要です。)

マッチングがより容易になるのは -

- 1.3.2.1 …あるデータ項目あるいはその項目の組み合わせによって、ある小さな、そして容易に識別可能な人の集団を隔離させるような場合。非常に小さな人口区分を示すコードを含めることは避けるべきです。例えば、極めて特定の地理との組み合わせたインディアンの種族又は詳細な職業。普通、1つのファイル又はリスト上に変数のグループが同時に現れることでもない限り、人は一度に1つ以上の変数を考える必要はありません。例えば、年齢と性別は外部ファイル上に同時に現れることはありそうですが、出生した国と職業はありそうにない;このように、ロシア生まれの建築家のような稀にしか出現しない事象を保護する必要はないとすべきです。

ファイルに含めるために申請したデータ項目のうち、小さく、容易に識別可能な人の集団を隔離させるような項目があれば、すべてリストアップして下さい。

この理由により、変更(すなわち、削除、再コード化、誤差(ノイズ)の付加)されるであろうデータ項目をすべてリストアップして下さい。

-
-
- 1.3.2.2 …当該ファイルが人口の実質的な一部を含んでいる(仮に $p > 0.5$ とする)場合。例:大雇用主、高収入の個人、様々なタイプの医者や科学者、または様々なタイプの施設の入所者。追加的な副標本(subsampling)では、データ提供前のある層内においてしばしば必要となります。

このような人の集団がファイル上にいくつか存在している、また、何らかの方法でこれらが副標本となっている場合、これらの人の集団を明らかにして下さい。

-
-
- 1.3.2.3 …当該ファイルに含まれる何らかの情報が記録または他のソースから得られたものであり、その情報が個々の識別子または詳細な地理情報を持っている外部ファイルへのリンクとして働くかもしれないような場合。その例に含まれるものとしては、公益事業会社から燃料消費又は費用の記録;(10年)人口・住宅センサスから得られた近隣、土地またはRDの集計での特徴;政府機関からの福祉または社会保障データ;警察からの逮捕記録;年金と健康保険として従業員に与えられた給付金。

ファイルに含めるために申請したデータ項目のうち、回答者(調査対象者)からのインタビューからは得られないことのない項目があれば、すべてリストアップして下さい。

この理由で変更または削除したデータ項目をすべてリストアップして下さい。

- 1.3.2.4 … マッチングにしばしば用いられる、正確な誕生日、性別、人種のようなデータ項目がファイルに含まれる場合や、または正確な収入総額、不動産税や他の税金、政府支援の(事業)プログラムの開始日と終了日などの両ファイルを一致させる他の項目が含まれる場合。

もしある場合、これらのデータ項目をすべてリストアップして下さい。

この理由で変更または削除したデータ項目をすべてリストアップして下さい。

- 1.3.2.5 … 長期的な(longitudinal)データが収集されている場合。すなわち、同じ回答者/単位を対象にしたデータがいくつか異なる照会期間に収集されるような場合。第一に懸念すべき事として話すべきことは、時系列のデータ項目は潜在的に外部の記録とマッチングしやすいということである。例えば、所得税あるいは雇用記録。もしデータが一度ならず同じ回答者から収集されるのであれば、インタビューの頻度、一つの標本単位として残存するかもしれない期間、ある時期から次の時期までに標本単位として一致する公算に影響する要因を示して下さい。
-
-

- 1.3.2.6 … 極めて特定の地理情報がファイルに含まれている場合。例えば、州、SMSAsなど。(この地理情報は、人口統計表で示されるべきです。)

- 1.3.2.7 … このファイルを外部のデータとマッチングできないよう保護することに関して、以前に述べられなかった考慮すべき事をすべて記述して下さい。例えば、データにある非信頼性またはもともとのノイズ(誤差)など。
-
-

1.3.3 特徴が唯一である集団を明らかにするクロス表

- 1.3.3.1 唯一の特徴を持つ集団が明らかになるようなクロス表を作成されましたか?

もし「いいえ」ならば、1.4に跳んで下さい。

1.3.3.2 結果はどうでしたか。

1.3.3.3 これらの結果が基になった開示リスクがある場合、そのリスクを減少させるための追加の手段として、どのようなものが取られるでしょうか。

1.4. 誤差(ノイズ(Noise))

1.4.1 データに何らかの誤差(ノイズ)を加えられましたか。 _____
もし「いいえ」ならば、1.5に跳んで下さい。

1.4.2 どのような手順を使ってデータに誤差を加えましたか。どうかその手順(すなわち、影響のあるレコードは何パーセントか、誤差の分布など)の詳細を教えてください。

想定されるもの：乱数による誤差(random noise)
レコードの交換 (record swapping)
階級の交換(rank swapping)
ブランク (blank)又は補定(imputation)

1.4.3 元のファイルに戻って誤差を付与したデータとのマッチングを行うような試みはありましたか。

もし「いいえ」ならば、1.5に跳んで下さい。

1.4.4 どのように行いましたが、また、マッチングの成功率はどのくらいでしたか。

1.5. 編集済データ(回答者が提供したデータ値を変更したもの)と補定済データ(無回答のために作られたデータ値)は、それ自身組み込まれた誤差(ノイズ)となります。編集(エディティング)と補定(インピュテーション)のプロセスは、開示に対する保護となります。その値が分る場合は、どうかこのセクションの質問に答えて下さい。

1.5.1 編集したデータ項目を少なくとも1つ含むレコードの割合は何パーセントですか。 _____

1.5.2 全データ項目の何パーセントが編集されたのですか。 _____

1.5.3 補定したデータ項目を少なくとも1つ含むレコードの割合は何パーセントですか。 _____

1.5.4 全データ項目の何パーセントが補定されたのですか。 _____

1.6. 他の問題

- 1.6.1 全標本を含むか、あるいはすべての層の標本を極めて高い比率($p > 0.5$)で含むファイルは単に標本のsubsample(副標本)を含んでいるだけのファイルよりも開示に結びつきやすい傾向にある。例えば、ある人が特定の調査を受けたことが知られているような場合、すべての標本がそのファイルで入手可能と仮定すれば、その人の記録と一致するマイクロデータファイルが見つげられると推測できることとなります。

このようなファイルを含んでいますか

___ すべての標本

___ 標本のsubsample (副標本)(もしそうであれば、抽出率の範囲を明確に述べて下さい)

-
- 1.6.2 プロジェクト・マネジャーは、公開用ファイル(the public-use file)では省かれた詳細情報が含まれる内部バージョンのファイルから特別集計(special tabulations)を実施する場合、秘密保護の問題が引き起こされる可能性があることを知っておくべきです。例えば、公開用ファイルで含まれず、ファイル上の複数のデータ項目によってクロス表にされるような特定の地理情報が提供されてしまう可能性があります。DRBは、これらの作表を再検討するための手順の概略を記述したガイドラインを用意しました。どうか、このガイドライン(「公開用マイクロデータの利用によってもたらされる調査作表の開示可能性("Disclosure Potential of Survey Tabulations Given the Availability of Public-Use Microdata")」)を参照して下さい。そして、もし公開用ファイルで入手できない詳細を用いて統計表を発表する計画があれば、どうかDRBに相談して下さい。

1.6.3 標本設計について手短かに記述して下さい

- 1)層化、クラスタリング、階層に関する何らかの記述を含めて下さい。そして標本となるある種の単位がどの段階においても $p > 0.5$ となることが明らかになることを含めて;
- 2)申請のあった標本単位、調査区単位、研究分析の単位との比較・対照を含めて下さい;
- 3)公的部門に提出予定あるいは予定がなくても標本設計(標本計画と推定量)の情報を明らかにして下さい。
- 4)ユーザーがどのように標本の分散を推定するか記述してください。申請されたファイルレイアウトで何らかの提供された「変数のネスト」や複製のアプローチのために使われる何らかのウエイトも潜在的に明らかにしながら。

1.6.4 補足

この情報は、別の調査への補足情報として集められたものですか _____

もし「いいえ」ならば、チェックリストのこのセクションは、完成です。

このマイクロデータファイルは、主要調査から作成されたファイルとリンクできますか。 _____

もし「はい」ならば、主要ファイルにはどのような地理情報がありますか。

セクション 2 人口統計の表形式のデータ

2.1. データ

2.1.1 この成果物(統計表のデータ)を「標準的でなく」しているものは何ですか(すなわち、標本人口が10万人未満の人口センサスデータあるいは識別可能な地理的区域)

2.1.2 これは標本データですかセンサスデータですか。 _____
もしセンサスであれば、2.1.5への跳んで下さい。

2.1.3 抽出率を含め、標本設計について手短かに述べて下さい。

2.1.4 例えば、あるセルが10と表示されていれば、当該調査においてたった1人が当該セルの特定の特性を持たされていることを意味するというように、ウエイトは共通の知識(あるいは簡単に推量されうるもの)となっていますか。

2.1.5 発表される表の記載内容(すなわち次元、変数とそのカテゴリー)について教えて下さい。

2.1.6 発表される地理的なレベルはどの程度ですか。

2.2. これらの表を作成するために使用された行政記録のデータはありますか。 _____
もし「はい」ならば、どうか記述して下さい。

2.3. 編集済データ(回答者によって提供されたデータの値が変更されたもの)と補定データ(無回答のために作られたデータの値)は、それ自身に組み込まれた「誤差(ノイズ)」を有しません。加工と補定のプロセスは、開示に対する保護となります。もし、その値がわかるのであれば、このセクションの質問に答えて下さい。

2.3.1 編集したデータ項目を少なくとも1つ含むレコードの割合は何パーセントですか。 _____

- 2.3.2 全データ項目の何パーセントが編集されたのですか。 _____
- 2.3.3 補定したデータ項目を少なくとも1つ含むレコードの割合は何パーセントですか。 _____
- 2.3.4 全データ項目の何パーセントが補定されたのですか。 _____

2.4. 開示回避

このデータに対してどのような開示回避のテクニックを利用したか、また、なぜ使用したか、(もしあるなら)どうか詳細な情報を提供して下さい。

想定される技術：レコードの交換 (record swapping)
ブランク (blank)又は補定(imputation)
階級の交換(rank swapping)
乱数による誤差(random noise)
セルの秘匿(cell suppression)
数値の丸めによる制御(controlled rounding)

セクション3 事業所統計の表形式のデータ

3.1. データ

3.1.1 これは標本データですかセンサスデータですか。 _____

3.1.2 事業所数は公表されますか。 _____
もし、センサスであれば3.1.5へ跳んで下さい。

3.1.3 あるタイプの事業所は、確信をもって選ばれましたか。 _____

3.1.4 抽出率を含む標本設計について手短かに記述して下さい。

3.1.5 発表されるデータは何ですか、また、形式(すなわち表の次元、変数とその詳細)はどうなっていますか。

3.1.6 発表される地理的なレベルはどの程度ですか。

3.2. これらの表を作成するために使用された行政記録のデータはありますか。 _____
もし「はい」ならば、どうか記述して下さい。

3.3. 編集済データ(回答者によって提供されたデータ項目の値が変更されたもの)と補定データ(無回答のために作られたデータの値)は、それ自身に組み込まれた「誤差(ノイズ)」を有します。加工と補定のプロセスは、開示に対する保護となります。もし、その値がわかるのであれば、このセクションの質問に答えて下さい。

3.3.1 編集したデータ項目を少なくとも1つ含むレコードの割合は何パーセントですか。 _____

3.3.2 全データ項目の何パーセントが編集されたのですか。 _____

3.3.3 補定したデータ項目を少なくとも1つ含むレコードの割合は何パーセントですか。 _____

3.3.4 全データ項目の何パーセントが補定されたのですか。 _____

3.4. 開示回避

このデータに対してどのような開示回避のテクニックを利用したか、また、なぜ使用したか、(もしあるなら)どうか詳細な情報を提供して下さい。

想定される技術：セルの秘匿(cell suppression)

誤差 (noise)

もし、セルの秘匿が使われなかったら、3.4.2に跳んで下さい。

3.4.1 セルの秘匿

3.4.1.2 一次秘匿の決定のため、どのような基準(また、どのようなパラメータで)を使用しましたか。 _____

もし、センサスのデータであれば、3.4.1.4へ跳んで下さい。

3.4.1.3 当該調査に適合させるため、その基準はどのように採用されたのですか？

3.4.1.4 一次秘匿を施すため、総数に何らかの保護をかけましたか？

3.4.1.5 一次秘匿の決定と各一次秘匿に必要な総数の保護に先立って、企業ごとに事業所を(又は所有者ごとに農場を)結合(名寄せ)しましたか。 _____

3.4.1.6 セルの秘匿を実施する際に、キーとなる項目を選びましたか。もしそうなら、それは何でしたか、又なぜですか。

3.4.1.7 セルの秘匿の実施は、センサス局の秘匿ソフトウェアによるものですか、又は手作業によるものですか。 _____

秘匿ソフトウェアによる場合は、3.4.1.9に跳んで下さい。

3.4.1.8 表の秘匿パターンは検査しましたか？ _____

もし、秘匿が手作業で実施された場合、3.4.2に跳んで下さい。

3.4.1.9 (あるかどうかわかりませんが)何らかのショートカットを使ったか、何らかのパラメータを用いましたか。

3.4.1.10 秘匿を手作業で除去したものはありますか。 _____
もしそうなら、なぜですか。

もしそうなら、それを置き換えるため、他のものをどのように選ばれるのですか。

3.4.1.11 三次元の表もすべて検査しましたか。

3.4.1.12 秘匿される値(すなわち、範囲、中央値、推計値、丸めた値、誤差を持たせた値など)として、何らかの追加情報も公表される予定ですか。

もしそうなら、どうか詳細について教えて下さい。

もし誤差が開示回避の技術として使われなかったら、3.4.3に跳んで下さい。

3.4.2 誤差(ノイズ(Noise))

3.4.2.1 どの項目に誤差を持たせましたか。

3.4.2.2 データへの誤差はどのように加えられましたか。

3.4.2.3 いくつの誤差がデータに加えられましたか。

3.4.3 このデータに何らかの他の開示回避技術が使われましたか。 _____
もしそうなら、どうか詳細について記述して下さい。

3.5. 特別なタイプのデータの取扱い

- 3.5.1 開示回避技術の適用によっては、特別な扱いを必要とするデータもあります。想定されるものとしては:

負の値となるデータ
ネット(正味)値でパーセントに変換したデータ
絶対値データ間の差
加重平均のデータ

特別な扱いを必要としたデータは、何かありましたか。_____

もしそうなら、どのデータで、どのように行いましたか。

3.6. 開示回避の調整

- 3.6.1 これは特別集計(special tabulation)ですか。_____

もし、「いいえ」ならば、3.6.3に跳んで下さい。

- 3.6.2 複数の秘匿パターンが相互に解き明かされないことを保証するため、すべての秘匿は、同じデータ・セットから生み出したすべての表の中で調整されなければなりません。

開示回避技術(例えばセル秘匿パターン)は、以前発表された基本表に用いられたものと調整されましたか。

- 3.6.3 同一(あるいはとてもよく似た)データは、別の課又は係によっても発表されたことがありますか。_____

もし「いいえ」ならば、チェックリストのこのセクションは完成です。

- 3.6.4 開示回避技術(例えばセル秘匿パターン)は、別の課又は係によって使われたものと調整されましたか。
