

諮問第 44 号の答申

国勢調査に係る匿名データの作成について（案）

本委員会は、総務省が作成を予定している国勢調査（以下「本調査」という。）に係る匿名データの作成方法の計画について審議した結果、下記の結論を得たので答申する。

記

1 計画の適否

本計画については、これにより作成される匿名データにおいて、本調査の調査客体の匿名性及び学術研究等における有用性がおおむね確保されるものと認められることから、適当である。

ただし、以下の「2 理由等」で指摘した事項については、計画の修正が必要である。

2 理由等

本調査は、我が国の人口の状況を明らかにするため、本邦内に常住している者を対象にほぼ5年ごとに実施している統計調査であり、10年ごとの大規模調査とその中間年の簡易調査とに大別される。今回、総務省は、大規模調査年と簡易調査年とで調査項目が異なること及び時系列での比較を考慮して、大規模調査である平成12年調査と簡易調査である平成17年調査の2か年分の匿名データを作成することを計画している。

また、その作成に当たって、これまで用いられてきた、情報の削除、一定の値を上限値としそれを上回る場合に上限値以上でまとめる措置（以下「トップコーディング」という。）や、分類の程度を粗いものにする措置（以下「リコーディング」という。）等の匿名化措置以外に、攪乱的な秘匿手法の一つである調査票情報のレコードの一部を別のレコードのものと入れ替える措置（以下「スワッピング」という。）を新たに導入する計画である。

これらの匿名データの具体的な作成方法に関する計画の適否等については、以下のとおりである。

（1）情報の削除

ア 地域区分

地域区分については、「都道府県」及び「人口50万以上の市区」とし、人口50万未満の市区町村コードは削除する計画である。

このうち、「都道府県」及び「人口50万以上の市区」を提供することについては、本調査が悉皆調査であり、地域レベルでの分析が可能であることが特徴の一つであるため、地域分析に対するニーズは高いと考えられること、本調査の結果では従来から人口50万以上の市区について詳細な統計表が提供されていることから、適当である。

また、人口50万未満の市区町村コードを削除することについては、人口50万未満の市区町村が明らかになると、外観識別性の高い属性情報と組み合わせた場合に調査客体が特定される可能性が生じることから、適当である。

イ レコードのサンプリング

本調査の匿名データの作成に当たっては、全ての調査票情報のレコードから、全世帯を母集団として1%を抽出することとし、世帯の種類（「一般世帯」及び「施設等の世帯」）ごとに、市区町村及び世帯人員等で並べ替えた上で、「一般世帯」については世帯単位で、また「施設等の世帯」については個人単位で抽出処理を行い、これらを統合したものをを用いる計画である。また、抽出された世帯は、乱数により並べ替える計画である。

これについては、次の理由から適当である。

- ① サンプリングは、匿名データの中に特定の調査客体が含まれるか否かの判断を困難とする措置であること
- ② サンプリングに当たっては、各市区町村における世帯数や各世帯の世帯人員等による分布を反映した抽出処理が行われるとともに、その抽出率は、都道府県のみならず、人口50万以上の市区など一定の人口規模以上の市区についての集計・分析が可能なデータ量が確保できるものとなっていること
- ③ 当該方法により作成された匿名データによる統計と既に公表されている本調査の全レコードから作成された統計（以下「公表統計」という。）との間で、男女別、年齢5歳階級別、世帯主との続き柄別、配偶関係別等でみた分布に大きな乖離は無く、当該匿名データの有用性が確保されていること

ただし、世帯の種類によって抽出単位が異なることから、分析等において支障が生じないように利用者に対して利用上の注意を示す必要がある。また、主要な項目に関して、匿名データによる統計と公表統計の結果を比較できるような情報を提供するなど、利用者の利便性の向上を図る必要がある。

ウ 直接的な識別情報の削除

本調査のレコードに含まれる情報のうち、調査区番号、世帯番号・調査区内連番など、調査客体が直接識別できる情報を削除する計画である。

これらについては、調査客体の特定や探索を防止するために効果的な措置であることから、適当である。

エ 出現頻度が低い又は特徴的な値があるレコードを含む世帯の削除

本調査のレコードのうち、表1に示す出現頻度が低い又は特徴的な値があるレコードを含む世帯を削除する計画である。

表1

対象となる世帯	具体的な計画の内容
世帯人員が多い世帯	地域区分ごとの出現頻度により、世帯人員が7人以上～9人以上いる世帯を削除
父子世帯	未婚、死別又は離別の男親と、その未婚の20歳未満の子供のみからなる世帯を削除
年齢差の大きい夫婦のいる世帯	年齢差が25歳以上の夫婦のいる世帯を削除
年齢差の大きい又は小さい親子のいる世帯	年齢差が45歳以上の親と子、年齢差が14歳以下の親と長子、又は19歳以下の親と末子のいる世帯を削除
世帯主又は配偶者のいずれか一方若しくは双方が外国人で子供の数が多い世帯	地域区分ごとの出現頻度により、子供の数が3人以上～7人以上の外国人世帯を削除

これらについては、次の理由から基本的に適当である。

- ① 世帯員数、性別、年齢等は、世帯の外部から比較的容易に把握可能な属性であり、トップコーディングやリコーディング等の匿名化措置を行ったとしても、地域情報等の他の属性情報と組み合わせた場合に出現頻度が低くなる世帯構成の場合は、調査客体が特定される可能性が生じること
- ② 地域によって出現頻度に大きな違いが見られる「世帯人員が多い世帯」や「世帯主又は配偶者のいずれか一方若しくは双方が外国人で子供の数が多い世帯」については、有用性の観点から地域区分ごとに削除基準を変える措置が講じられていること
- ③ 複数の変数の組合せをみる必要がある「年齢差の大きい夫婦のいる世帯」、「年齢差の大きい又は小さい親子のいる世帯」、「世帯主又は配偶者のいずれか一方若しくは双方が外国人で子供の数が多い世帯」については、実数値による分布に基づいて削除数を減らす措置が講じられていること

ただし、年齢差が45歳以上の親と子を削除することについては、親の性別によって分布が大きく異なっていることから、親の性別によって対象とする年齢差を変えることとし、男親については、年齢差を55歳以上に引き上げる必要がある。

オ 公表統計により母集団一意又は二意であることが判明しているレコードを含む世帯の削除

公表統計により、全国において、個体が識別されやすい属性情報を組み合わせた場合に、母集団の中で個体が1つしかないような母集団一意の状態であるか、又は同じ組合せを持つ個体が2つしかないような母集団二意の状態であることが判明しているレコードが含まれる世帯については、これを削除する計画である。加えて、公表統計のうち、特に外観識別性が高いと考えられる項目が含まれている統計表については、地域（「都道府県」及び「人口50万人以上の市区」）において、母集団一意又は二意であることが判明しているレコードが含まれる世帯も削除する計画である。

これらについては、国勢調査が悉皆調査であることや、詳細な集計表が外部参照情報として利用可能であることなどの特性を踏まえて、全国において母集団一意又は二意である場合や、

地域において、特に外観識別性が高いと考えられる項目が含まれている統計表で母集団一意又は二意である場合は、調査客体が特定される可能性が生じることから、適当である。

(2) 識別情報の分類区分の再編等

ア 世帯員に関する項目の再編等

(ア) 年齢

世帯員の年齢については、85歳以上をトップコーディングすることとし、トップコーディングを行う高齢者を除く0～84歳をリコーディングして5歳階級別とする計画である。

これらについては、各歳別の年齢や出現頻度の低い高齢者の年齢が明らかになると、今回提供される詳細な地域情報や他の外観識別性の高い属性情報との組み合わせにより調査客体が特定される可能性が生じることから、やむを得ない措置である。

(イ) 世帯主との続き柄、国籍、5年前の住居の所在地等

世帯主との続き柄、国籍、5年前の住居の所在地（平成12年国勢調査）、従業上の地位、職業（大分類）及び常住地による従業地・通学地については、それぞれ表2に示す分類区分の再編等を行う計画である。

表2

項目	具体的な計画の内容
世帯主との続き柄	「他の親族」、「住み込みの雇人」及び「その他」を統合
国籍	「日本人」及び「外国人」の2区分とし、外国籍の内訳は提供しない
5年前の住居の所在地 (平成12年国勢調査)	「他県から」及び「国外から」を統合
従業上の地位	「雇人のある業主」、「雇人のない業主」及び「家庭内職者」を統合
職業（大分類）	「保安職業従事者」、「農林漁業作業員」及び「運輸・通信従事者」を統合
常住地による従業地・ 通学地	「県内他市区町村で従業・通学」及び「他県で従業・通学」を統合

これらについては、出現頻度の低い分類区分を統合するか又は提供しないことによつて、地域情報や他の外観識別性の高い属性情報との組み合わせによる調査客体の特定を防ぐことから、適当である。

なお、5年前の住居の所在地における「他県から」の場合の調査項目である5年前の常住都道府県等については、「国外から」との統合のため提供されないこととなるが、これは統合に際して「国外から」の匿名性を確保する必要があることから、適当である。

(ウ) 労働力状態

労働力状態については、「家事などのほか仕事」及び「通学のかたわら仕事」を統合する

計画である。

これについては、「通学のかたわら仕事」は、地域情報や他の属性情報と組み合わせた場
合に出現頻度が低くなり、調査客体を特定される可能性が生じることから、簡易調査である
平成 17 年調査の場合は、計画のとおり統合することが適当であるが、大規模調査である平成
12 年調査の場合は、提供される調査項目が多く、「家事などのほか仕事」と統合するのみで
は秘匿が不十分となることから、就業者の内訳を提供しないように変更することにより、匿
名性を確保する必要がある。

(エ) 就業時間

就業時間については、実数及び公表統計と同じ階級（14 区分及び 10 区分）で提供するこ
ととし、実数については 90 時間以上をトップコーディングする計画である。

このうち、出現頻度が低い長時間労働をトップコーディングすることについては、これに
より、他の属性情報との組み合わせによる調査客体の特定を防ぐことから、適当である。た
だし、トップコーディングを行った当該変数については、その地域区分別の平均値等を匿名
データの提供に併せて提供することにより、利用者の利便性向上を図る必要がある。

また、公表統計と同じ階級で提供することについては、公表統計との整合性を図ること
により匿名データの利便性が向上することから、適当である。

(オ) 産業（大分類）

産業については、大分類で提供することとし、以下の分類区分の統合を行う計画である。

- ① 「農業」、「林業」及び「漁業」を統合
- ② 「鉱業」及び「建設業」を統合
- ③ 「製造業」及び「電気・ガス・熱供給・水道業」を統合
- ④ 「複合サービス業」及び「サービス業（他に分類されないもの）」を統合

これらについては、「農業」、「林業」、「漁業」、「鉱業」、「電気・ガス・熱供給・
水道業」及び「複合サービス業」は、地域別にみると出現頻度が低くなる場合があり、他の
属性情報との組み合わせによって調査客体を特定される可能性が生じること、統合に当たっ
ては、できるだけ類似性のある項目を統合する方がよいと考えられることから、基本的に適
当である。

ただし、「農業」、「林業」、「漁業」については、これらを統合したとしても秘匿が不
十分な場合があることから、特定化の危険性が高いレコードを含む世帯を削除することによ
り、匿名性を確保する必要がある。

(カ) 利用交通手段（平成 12 年国勢調査）

利用交通手段については、「利用交通手段が 1 種類」に関して、「勤め先・学校のバス」、
「ハイヤー・タクシー」、「オートバイ」及び「その他」を統合する計画である。

これについては、利用交通手段が 1 種類の場合は外観識別性が高いと考えられる一方、複
数の場合には 2 種類目以降の交通手段を識別することが難しいこと、「勤め先・学校のバス」、
「ハイヤー・タクシー」及び「オートバイ」は、地域情報や他の属性情報と組み合わせた場

合に出現頻度が低くなり、調査客体を特定される可能性が高くなることから、基本的に統合による匿名性の確保が必要である。

ただし、統合に当たっては、できるだけ交通手段として類似性のあるものを統合する方が望ましく、「オートバイ」と「自転車」に関しては、産業や職業別の就業者割合等をみても、分布状況が似ていることから、「オートバイ」及び「自転車」を統合することとし、残りの「勤め先・学校のバス」、「ハイヤー・タクシー」及び「その他」を統合するように変更する必要がある。

イ 世帯に関する項目の再編等

(ア) 世帯の種類、世帯人員、世帯の家族類型等

世帯の種類、世帯人員、世帯の家族類型、住居の種類・住宅の所有の関係、住宅の建て方については、それぞれ表3に示す分類区分の再編等を行う計画である。

表3

項目	具体的な計画の内容
世帯の種類	「一般世帯」及び「施設等の世帯」の2区分を提供することとし、「施設等の世帯」の内訳は提供しない
世帯人員	「施設等の世帯」については提供しない
世帯の家族類型	公表統計と同じ6区分により提供
住居の種類・住宅の所有の関係	「住宅に住む一般世帯」の以下の分類区分を統合することとし、「住宅以外に住む一般世帯」の内訳（「会社等の独身寮」及び「その他」）は提供しない ①「公営の借家」及び「都市機構・公社の借家」を統合 ②「給与住宅」及び「間借り」を統合
住宅の建て方	「長屋建」及び「その他」を統合

これらについては、出現頻度の低い分類区分を統合するか又は提供しないことによって、地域情報や他の外観識別性の高い属性情報との組み合わせによる調査客体の特定を防ぐこと、世帯の家族類型は製表の過程で作成されたものであるが、国勢調査の特徴を表す分類の一つであり利用ニーズは高いと考えられ、公表統計との整合性を図ることにより匿名データの利便性が向上することから、適当である。

(イ) 家計の収入の種類（平成12年国勢調査）

家計の収入の種類については、これを提供しない計画であるが、家計の収入の種類は、調査に際し世帯の忌避感が大きい調査項目であり、今後の国勢調査の実施に対する影響を懸念する指摘はあるものの、利用者にとっては有益な情報であり、出現頻度の低い分類区分を統合することにより匿名性を確保することが可能なことから、以下のような匿名化措置を講じた上で提供するように計画を変更する必要がある。

- ① 「賃金・給料が主な世帯」のうち、「農業収入もある世帯」及び「その他」を統合
- ② 「農業収入が主な世帯」及び「農業収入以外の事業収入が主な世帯」を統合し、その

内訳は提供しない

- ③ 「内職収入が主な世帯」及び「その他の収入が主な世帯」を統合し、その内訳は提供しない

(ウ) 住宅の床面積、建物全体の階数及び世帯が住んでいる階

住宅の床面積、建物全体の階数及び世帯が住んでいる階については、それぞれ表4に示す分類区分の再編を行う計画である。

表4

項目	具体的な計画の内容
住宅の床面積	実数ではなく、基本的に公表統計に合わせた階級で提供することとし、この際、「200～249 m ² 」及び「250 m ² 以上」を統合
建物全体の階数	実数ではなく、基本的に公表統計に合わせた階級で提供することとし、地域区分ごとの出現頻度により、必要に応じて、「6～10階建」、「11～14階建」、「15階建以上」の区分のうち、「6～10階建」以上、あるいは「11～14階建」以上の区分を統合
世帯が住んでいる階	実数ではなく、基本的に公表統計に合わせた階級で提供することとし、地域区分ごとの出現頻度により、必要に応じて、「3～5階」、「6～10階」、「11～14階」、「15階以上」の区分のうち、「3～5階」以上、「6～10階」以上、あるいは「11～14階」以上の区分を統合

これらについては、住宅の床面積及び階数は外観識別性が高く、実数が明らかになると、今回提供される詳細な地域情報や他の外観識別性の高い属性情報との組み合わせにより調査客体が特定される可能性が生じること、出現頻度の低い高層階は地域による差が大きく、有用性の観点から地域区分ごとに統合する階数区分を変える措置が講じられていることから、適当である。

(3) その他の匿名化措置

ア スワッピング

本調査の匿名データの作成に当たっては、新たな匿名化措置としてスワッピングを行うこととし、一部世帯を他の地域の類似世帯と入れ替える計画である。

このスワッピングの導入及び方法については、次の理由から適当である。

- ① 今回、地域区分として「都道府県」及び「人口50万以上の市区」を提供することとしており、地域情報が強力な識別情報となり得ること、また、国勢調査は悉皆調査であることから、既に公表されている詳細な集計表自体が有力な外部参照情報として利用可能であり、これらから想定される範囲で情報の削除やリコーディング等の措置を講じることに加え、更に攪乱的な秘匿手法を用いることにより、属性情報を様々に組み合わせた場合においても匿名性が確保されるように措置することは必要であると考えられること
- ② 本計画の方法によるスワッピングの前後で、影響を受ける項目の分布にあまり差異はな

く、有用性を損なう程のものではないと考えられること

ただし、利用者にとって、今回のスワッピングによる影響の範囲がどの地域までかは極めて重要である一方、スワッピングは同一都道府県内で行う計画であることを開示したとしても、匿名性は確保されていると考えられることから、利用者に対してこれを開示することにより、本匿名データの有用性を高める必要がある。

イ 匿名データの提供時期

本計画では、匿名データの作成対象調査を調査実施後5年以上経過したものとしている。

これについては、有用性の観点からは、最新のデータに基づく匿名データの作成に対する強いニーズがあり、匿名データの提供時期を早められないかという指摘はあるものの、匿名性の観点からは、調査実施後5年以上経過することで属性の変化により個人を特定し難くする効果があること、及び本調査が本邦内に常住している者を対象とした悉皆調査であり、実査への影響も考慮した場合に直近の調査の匿名データは提供し難いことから、本調査に関する匿名データの提供時期の短縮化は困難であり、やむを得ない措置である。

3 今後の課題

(1) トップコーディングを行う高齢者の年齢の検討

本計画では、世帯員の年齢については、85歳以上でトップコーディングを行うこととしている。しかし、トップコーディングの上限値については、近年の急速な高齢化の進展や高齢者に関する分析の重要性等を踏まえ、今後、匿名データの作成対象年次を拡大する際には、当該年次の人口の年齢構成に応じた検討が必要である。

(2) 複数の匿名データの作成の可能性に関する検討

今回の匿名データの作成に当たっては、悉皆調査の調査票情報を世帯単位で、また、地域区分を「都道府県」及び「人口50万以上の市区」として提供することを前提に考えられていることから、情報の削除や識別情報の分類区分の再編等において、厳格な匿名化措置を講じていることはやむを得ない。

しかし、匿名データの利用者のニーズについては様々なものが考えられ、例えば、世帯員の年齢を各歳別とすること、あるいは世帯の削除や産業等の分類区分の統合を緩和することなどに対するニーズが指摘されている。これについては、地域区分を全国のみとすることや提供データを個人単位とすることなどにより、匿名化措置を緩和できる可能性も考えられることから、今後、利用者のニーズを踏まえて、匿名化措置の内容や組合せ、抽出単位の異なる複数の匿名データの作成の可能性について検討する必要がある。

第13回 匿名データ部会 議事概要

1 日 時 平成25年1月31日(木) 9:58~11:58

2 場 所 中央合同庁舎4号館2階 共用第3特別会議室

3 出席者

(部会長) 椿広計

(委員) 安部由起子、津谷典子

(専門委員) 伊藤伸介、加藤久和、安田聖

(審議協力者) 総務省(政策統括官(統計基準担当))、財務省、厚生労働省、農林水産省、経済産業省、国土交通省、東京都

(諮問者) 総務省統計局統計調査部: 井上調査企画課長、岩佐国勢統計課長ほか

(事務局) 内閣府統計委員会担当室: 若林参事官ほか

4 議 事

(1) 国勢調査に係る匿名データの作成について

(2) その他

5 議事概要

(1) 前回部会での質問に対する対応、及び残りの論点について

ア 年齢差の大きい親子のいる世帯の削除の状況について

前回部会において、有用性の観点から、年齢差が45歳以上の親と子のいる世帯について削除の状況を確認したいという意見があったことに対して、諮問者から、特に男親の場合に削除数が非常に多くなることから、親の性別による分布の差を考慮して、男親について年齢差を55歳以上とするという提案があった。各委員等からの主な意見は次のとおり。

- ・ 年齢差の大きい親子のいる世帯の削除数が多くなることを心配していた。この変数については、親の性別によって削除する年齢差を変える方法でよいと思う。

《部会長のまとめ》

- ・ 年齢差が大きい親子のいる世帯について、親の性別によって削除する年齢差を変えることとし、男親の場合に年齢差を55歳以上とすることは、これにより削除数を減らすことができるので、認めることとしたい。

イ 利用交通手段について

前回部会において、「利用交通手段」について、「オートバイ」と「自転車」は共に二輪で性質が似ていると考えられるため、これらを統合することはできないかという意見があったことに対して、諮問者から、「オートバイ」及び「自転車」を統合し、残りの「勤め先・学校のバス」、「ハイヤー・タクシー」及び「その他」を統合することとしたとしても、秘匿に関して特段問題は生じないという説明があった。各委員等からの主な意見は次のとおり。

- ・ 「オートバイ」及び「自転車」は性質が似ているので、秘匿上の問題がないのであれば、利用者の立場から見た場合に、これらを統合するのが望ましいと思う。

《部会長のまとめ》

- ・ 有用性の観点から、「オートバイ」及び「自転車」を統合し、残りの「勤め先・学校のバス」、「ハイヤー・タクシー」及び「その他」を統合することは、秘匿に関して特段問題は生じないと考えられるので、認めることとしたい。

ウ トップコーディングが行われた変数の基本統計量の提供について

前回部会において、トップコーディングが行われた変数の基本統計量の提供について地域ごとに提供できないかという意見があったことに対して、諮問者から、対象となる就業時間について地域ごとに提供するという説明が行われ、了承された。

エ 労働力状態について

- ・ 「労働力状態」については、「通学のかたわら仕事」を秘匿するため、「家事などのほか仕事」と統合することとしているが、平成 12 年調査で提供される調査項目と組み合わせた場合に、匿名性に問題は生じないか。
- 当初、「家事などのほか仕事」と「通学のかたわら仕事」を統合することを考えていたが、平成 12 年調査については、当該調査に含まれる調査項目との組み合わせによっては秘匿が不十分となる場合があることから、①その調査項目自体を提供しないこととするか、②就業者の内訳を提供しないこととするか、③更に「休業者」を加えた3つを統合することとするか、あるいは、④統合の仕方は変えず、外観識別性が高いと考えられるレコードを削除することとするかのいずれかにより対応したい。
- ・ 平成 12 年調査にある調査項目は有用性が高いこと、また削除数を最小限にとどめる方がよいということから、分類区分の統合か、あるいは就業者の内訳を提供しないこととする方がよい。
- ・ 「家事などのほか仕事」、「通学のかたわら仕事」に「休業者」を統合すれば問題は起こらないのではないか。「労働力状態」については自己申告なので、通学と仕事のどちらが主であるかもわかりにくいと考えられる。
- ・ 分類区分の統合による対応では、地域情報や他の情報と組み合わせた場合にまだ「通学のかたわら仕事」の秘匿が不十分となる可能性があるため、「通学のかたわら仕事」の出現頻度が低くなる地域を対象に外観識別性が高いと思われるレコードを削除する方がよい。
- ・ 外観識別性が高いと思われるレコードの削除という例外的な処理をするべきではなく、一定の方法で扱った方がよいのではないか。
- ・ 提供可能な調査項目を考えて、その上で 0.5%基準に照らして分類区分の統合を行い、それでも秘匿が難しいという場合にレコードの削除を行うということであり、その意味では、例外にはならないのではないかと考える。
- ・ 更に「休業者」を統合するだけでは秘匿性の観点でリスクがあるということであれば、有用性は劣るものの、就業者の内訳を提供しないとするのがよいのではないか。
- ・ 今回は就業者の内訳を提供しないことに賛成するが、今後検討の余地がある。

《部会長のまとめ》

- ・ 秘匿を重視する観点から、平成 12 年調査については就業者の内訳を提供しないということでもとめることとしたい。

オ 「スワッピング」の説明について

- ・ 「スワッピング」については、その具体的な方法を一般に説明することができないことから、例えば、スワッピングの前後で影響を受ける項目の分布にあまり差異がないというようなことを、ユ

一ザがわかるような形で説明するということが考えられるが、総務省統計局ではどのようなことを考えているか。

- スワッピングを行っていること、及びそれによる誤差は標本誤差と比較して許容範囲となっており有用性に支障はないことについて、ホームページ等で説明することを考えている。また、主要な項目の全国の結果について、母集団と匿名データとを比較した場合にそれほど乖離がないという情報についても、併せて提供したいと考えている。
 - ・ 母集団と匿名データとの比較の情報について、全国のみ提供するということが、特定の地域のデータを分析したいと考えている利用者にとっては、それでは情報が足りないのではないか。
- 利用者の関心のある地域に応じてそのようなデータを個別に提供することは困難であり、提供するデータの量も多くなってしまふ。地域別の結果の比較については、匿名データによる集計結果と詳細な地域の公表統計により行うことができる。
 - ・ 国勢調査では地域別の詳細な報告書が公表されているので、スワッピングで基本的な属性に歪みは生じないということを示すという点では、全国の結果の比較だけを提供すればよいのではないか。
 - ・ 学会や学術誌での発表などを行う際に、どの程度の詳細な地域までの分析が保証されるのかが、利用者が一番の関心事項であると思うが、今回の説明では、その点がわからない。例えば、都道府県の範囲の分析であれば保証されている、あるいは都道府県内でスワッピングを行っているといったようなことを示すことはできないか。
 - ・ 地域での分析が保証されているということを説明できれば、査読付き学術誌への投稿論文の審査等などの関係で、利用価値は非常に大きくなると考えられる。
 - ・ 利用者がスワッピングについて詳しいとは限らないので、諸外国の人口センサスの匿名化の際にも用いられている手法であるということなどを、ウェブサイト等に掲載するという対応をしていたらよいと思う。
 - ・ 利用者への説明については、書きぶりにもよるが、スワッピングが都道府県内で行われていることを開示することのリスクは、基本的にあまりないと考えられる。

《部会長のまとめ》

- ・ 都道府県内では分析に影響を与えないということを説明することで、研究の有用性が格段に向上し、そのことによるリスクはそれほど増えていないと考えられるので、スワッピングが都道府県内で行われているということを開示すべきである。

(2) 答申(案)の審議について

椿部会長から、資料「諮問第44号の答申『国勢調査に係る匿名データの作成について』(案)」が示され、事務局からの答申(案)についての説明の後、審議が行われた結果、一部で所要の修正を行うこととされたが、答申(案)はおおむね適当であるとして採択された。なお、答申(案)の修正部分の表現については、部会長に一任することとされた。各委員等からの主な意見は次のとおり。

ア 出現頻度が低い又は特徴的な値があるレコードを含む世帯の削除(2の(1)のエ)

- ・ 本日の議論で変更のあった年齢差の大きい親子のいる世帯の削除については、特に父子について削除数が非常に多くなっていることから、年齢差が55歳以上の男親と子と修正する必要がある。

イ 労働力状態(2の(2)のアの(ウ))

- ・ 平成12年調査に関しては、秘匿が不十分となることから、就業者の内訳を提供しないようにする必要があるのであるという修正を行うこととする。

ウ 利用交通手段（２の（２）のアの（カ））

- ・ 「オートバイ」と「自転車」を統合するというので、答申案の案２を採用することとする。

エ スワッピング（２の（３）のア）

- ・ スワッピングの導入について、「やむを得ない」と記述している部分については、国勢調査は悉皆調査であり秘匿性の観点からスワッピングを行う必要があること、及び諸外国の人口センサスでも導入していることから、はっきりと「必要である」と表現とした方がよい。
- ・ スワッピングの導入が都道府県の分析に影響を与えないことについては、そのことを開示しても調査客体の特定化の危険性が低いということを明記した方がよい。

《部会長のまとめ》

- ・ スワッピングの導入について、「やむを得ない」となっている部分については、「必要である」と修正したい。
- ・ スワッピングが行われていることがわかっている以上、特定化のリスクはあまり増えないということを中心に、スワッピングは都道府県内で行われているということを開示するように修正したい。

オ トップコーディングを行う高齢者の年齢の検討（３の（１））

- ・ トップコーディングを行う高齢者の年齢の検討に関して、今後の作成対象年次の拡大の際に「人口構成」に応じた検討が必要と記述している部分については、どのような人口の構成かはっきりとわかるように、「人口の年齢構成」とした方がよい。

《部会長のまとめ》

- ・ 「人口構成」という表現については、「人口の年齢構成」に修正することとしたい。

カ 複数の匿名データの作成の可能性に関する検討（３の（２））

- ・ 「提供データを個人単位とすることなどにより」の「など」の中には、地域区分をより詳細にするという匿名化を緩和する方向だけでなく、場合によっては詳細な地域情報の提供を可能にするために、ある特定の調査項目を提供しないといった、匿名化を厳しくする方向も含まれていると考えた方がよい。
- ・ 匿名化措置については、その技法を中心にこれまで議論されているが、諸外国では、例えば学術目的でより詳細なデータを提供するために、法制度の観点から匿名データの提供方法を変えているところもあり、将来的には、そのような観点からの匿名データの提供の可能性もあり得るのではないか。

（３）その他

- ・ 本日は承された答申（案）、及び議事概要については、２月１５日（金）の統計委員会に提出することとされた。

以上

<文責 内閣府大臣官房統計委員会担当室 速報のため事後修正の可能性あり>



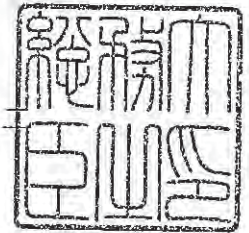
総統調第485号

平成24年10月26日

統計委員会委員長

樋口 美雄 殿

総務大臣
樽 床 伸



諮問第44号

国勢調査に係る匿名データの作成について（諮問）

標記について、別紙のとおり作成するに当たり、統計法（平成19年法律第53号）第35条第2項の規定に基づき、統計委員会の意見を求める。

諮問の概要

(国勢調査に係る匿名データの作成について)

今回、総務省は、国勢調査（平成 12 年及び 17 年）について、統計法（平成 19 年法律第 53 号）第 35 条第 1 項の規定に基づき、以下のとおり匿名データの作成を行う予定である。

1 国勢調査の匿名データを作成する理由

国勢調査は、我が国の人口の状況を明らかにするため、本邦内に常住している者を対象にほぼ 5 年ごとに実施している統計調査であり、学術研究や高等教育において利用ニーズが高い調査であることから、匿名データを作成するものである。

2 作成する匿名データの種類

「平成 12 年国勢調査（大規模調査）」及び「平成 17 年国勢調査（簡易調査）」とする。

大規模調査と簡易調査とで調査項目が異なること及び時系列での比較を考慮して、2 か年分を作成する。

3 匿名データの作成方法の概要

適用する匿名化措置は、以下のとおりである。

- ・国勢調査の全世帯数（母集団）から 1 %の割合で匿名データ用のファイルを作成する（サンプリング）。
- ・直接的な識別情報は、レコードから全面的に削除する。また、レコードの配列順が意味を成さないように、無作為に並べ替えを行う（識別情報の削除等）。
- ・発生頻度の低いレコード又は特徴的な値があるレコードを含む世帯を削除する（裾切りによるレコード削除）。
- ・既存の統計表により母集団一意又は二意であることが判明しているレコードを含む世帯を削除する（母集団一意等の削除）。
- ・極端に大きな値は、上限値を設けて頭打ちにする（トップコーディング）。
- ・分類区分の程度は、詳細なものではなく、粗いものとする（リコーディング）。
- ・一部世帯を他の地域の類似世帯と入れ替える（スワッピング）。

国勢調査（平成12年及び17年）の匿名データの作成方法（案）

国勢調査については、「匿名データの作成・提供に係るガイドライン」（平成 21 年 2 月 17 日付け総務省政策統括官（統計基準担当）決定）に示された基準を踏まえ、調査票情報（個票データ）に対して以下の処理を施すことにより匿名データを作成する。

なお、国勢調査の調査票情報は、1レコード1個人という編成であり、1レコードにはその個人が属する世帯の情報も含まれている。

1 作成する調査年次

「平成 12 年国勢調査（大規模調査）」及び「平成 17 年国勢調査（簡易調査）」とする。

なお、各年次における調査票情報のレコード数は次のとおり。

- ・平成 12 年：約 1 億 2 千 7 百万レコード（約 1 億 2 千 7 百万人、約 4 千 7 百万世帯）
- ・平成 17 年：約 1 億 2 千 8 百万レコード（約 1 億 2 千 8 百万人、約 5 千万世帯）

2 ファイルの種類

「世帯単位のファイル」を作成する。

3 作成に使用するデータ

全数データとする。

なお、使用する全数データには、調査票に記入された内容を数値や符号などに置き換えた電磁的記録を用いるため、調査票に文字で記入された「氏名」や「都道府県名・市区町村名」などは含まれていない。

4 地域区分

「都道府県」及び「人口50万以上の市区」とする。その理由は以下のとおり。

- (1) 地域区分に対するニーズが高いこと
- (2) 国勢調査結果（基幹統計）では、従来から人口50万以上の市区について、詳細な統計表を提供していること

5 サンプルング

母集団の 1% をサンプルングする。その理由は以下のとおり。

- (1) 母集団の 1% のデータ量は、都道府県のみならず、人口50万以上の市区など、一定の人口規模以上の市区についての集計・分析が可能と考えられること
- (2) 先行提供調査においても、総人口の 1% 以下としていること

調査名	リサンプルング率	総人口比率
全国消費実態調査	80%	0.1%
社会生活基本調査	80%	0.1%
就業構造基本調査	80%	0.6%
住宅・土地統計調査	10%	0.6%

なお、サンプルングされたデータに対し、以下の匿名化処理を施す。

6 サンプルング方法

世帯の種類(一般世帯及び施設等の世帯)ごとに抽出処理を行い、これらを統合して提供用匿名データとする。

- (1) 「一般世帯」については世帯単位で抽出する。「施設等の世帯」については個人単位で抽出する。
- (2) 「一般世帯」は市区町村コード>世帯人員>調査区番号>世帯番号の順に、「施設等の世帯」は市区町村コード>調査区番号>世帯番号の順にソートする。
- (3) 1から100までの数字の中から、無作為に選んだ数を抽出起番号とし、抽出間隔を100として、「一般世帯」は世帯単位、「施設等の世帯」は個人単位に、系統抽出法により抽出する。
- (4) 抽出された世帯は乱数を用いて並び替える。

7 情報の削除

- (1) 直接的な識別情報の削除

人口50万未満の市区町村コード、調査区番号、世帯番号・調査区内連番など、調査客体を直接識別できる情報を削除する。

- (2) 発生頻度の低い又は特徴的な値があるレコードの削除

発生頻度が低く、又は特徴的な値があり、分類区分の再編等の匿名化措置によっても、なお調査客体が特定される可能性の高い以下のレコードを含む世帯を削除する。

- ① 世帯人員が多い世帯

地域区分により、世帯人員が7～9人以上いる世帯を削除

- ② 父子世帯

- ③ 年齢差の大きい夫婦のいる世帯

年齢差が25歳以上の夫婦のいる世帯を削除

- ④ 年齢差の大きい又は小さい親子のいる世帯

年齢差が45歳以上の親と子、年齢差が14歳以下の親と長子又は19歳以下の親と末子のいる世帯を削除

- ⑤ 世帯主又は配偶者のいずれか一方若しくは双方が外国人で子供の数が多い世帯

地域区分により、子供の数が3～7人以上いる外国人世帯を削除

- (3) 既存の統計表により母集団一意又は二意であることが判明しているレコードを含む世帯の削除

既存の統計表により、全国において母集団一意又は二意であることが判明しているレコードが含まれる世帯を削除する。加えて、既存の統計表のうち、特に外観識別性が高いと考えられる項目が含まれる統計表については、地域(都道府県、人口50万以上の市区)において母集団一意又は二意であることが判明しているレコードが含まれる世帯も削除する。

8 分類区分の再編

提供する分類区分は、報告書の分類区分を基本とし、調査客体が特定される可能性の高い

分類区分については、他の分類区分との統合又はトップコーディングにより再編する。

(1) 世帯員に関する項目の再編

① 年齢

ア 0歳～84歳

5歳階級別で提供する。

イ 85歳以上

トップコーディングする。

② 世帯主との続柄

「他の親族」、「住み込みの雇人」及び「その他」を統合する。

③ 国籍

「日本人」及び「外国人」の2区分とし、外国籍の内訳は提供しない。

④ 5年前の住居の所在地（平成12年国勢調査）

「他県から」及び「国外から」を統合する。

⑤ 労働力状態

「家事などのほか仕事」及び「通学のかたわら仕事」を統合する。

⑥ 就業時間

ア 実数及び階級で提供する。

イ 実数については、90時間以上をトップコーディングする。

⑦ 従業上の地位

「雇人のある業主」、「雇人のない業主」及び「家庭内職者」を統合する。

⑧ 産業（大分類）

類似の分類区分を、以下のように統合する。

ア 「農業」、「林業」及び「漁業」を統合する。

イ 「鉱業」及び「建設業」を統合する。

ウ 「製造業」及び「電気・ガス・熱供給・水道業」を統合する。

エ 「複合サービス業」及び「サービス業（他に分類されないもの）」を統合する。（平成17年国勢調査）

⑨ 職業（大分類）

「保安職業従事者」、「農林漁業作業者」及び「運輸・通信従事者」を統合する。

⑩ 常住地による従業地・通学地

「県内他市区町村で従業・通学」及び「他県で従業・通学」を統合する。

⑪ 利用交通手段（平成12年国勢調査）

「利用交通手段が1種類」について、「勤め先・学校のバス」、「ハイヤー・タクシー」、「オートバイ」及び「その他」を統合する。

(2) 世帯に関する項目

① 世帯の種類

ア 「一般世帯」及び「施設等の世帯」の2区分を提供する。

イ 「施設等の世帯」の内訳は提供しない。

② 世帯人員

- ア 「一般世帯」について、地域区分により、世帯人員が7～9人以上いる世帯を削除。
- イ 「施設等の世帯」については提供しない。
- ③ 世帯の家族類型
 - 6 区分による提供
- ④ 家計の収入の種類（平成12年国勢調査）
 - 提供しない。
- ⑤ 住居の種類・住宅の所有の関係
 - ア 「公営の借家」及び「都市機構・公社の借家」を統合する。
 - イ 「給与住宅」及び「間借り」を統合する。
 - ウ 「住宅以外に住む一般世帯」の内訳（「会社等の独身寮」及び「その他」）は提供しない。
- ⑥ 住宅の床面積
 - ア 実数ではなく、階級で提供する。
 - イ 「200～249㎡」及び「250㎡以上」を統合する。
- ⑦ 住宅の建て方、建物全体の階数及び世帯が住んでいる階
 - ア 「長屋建」及び「その他」を統合する。
 - イ 共同住宅
 - （建物全体の階数）
 - 実数ではなく、階級で提供する。また、地域区分により、以下のとおりの匿名化措置を行う。
 - （ア） 匿名化措置なし
 - （イ） 「6～10階建」以上の区分を統合する。
 - （ウ） 「11～14階建」以上の区分を統合する。
 - （世帯が住んでいる階）
 - 実数ではなく、階級で提供する。また、地域区分により、以下のとおりの匿名化措置を行う。
 - （ア） 匿名化措置なし
 - （イ） 「3～5階建」以上の区分を統合する。
 - （ウ） 「6～10階建」以上の区分を統合する。
 - （エ） 「11～14階建」以上の区分を統合する。

9 スワッピング

地域情報は強力な識別情報となり得ることから、一部世帯を他の地域の類似世帯と入れ替える。

匿名データのチェックリスト（世帯調査用）
「匿名データの作成・提供に係るガイドライン」
（総務省政策統括官（統計基準担当）決定）準拠

匿名データを作成する統計データの名称及び年次

統計調査名：国勢調査
調査年：平成 12 年及び 17 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。

①地理情報のレベル：都道府県、人口 50 万以上の市区
②地理情報の加工の有無： 有 無
項目：市区町村コード、調査区番号、世帯番号・調査区内連番など
方法：人口 50 万未満の市区町村コード、調査区番号、世帯番号・調査区内連番などを削除するとともに、世帯単位にランダムに並べ替えを行う

(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。

地理情報以外の地理的情報：有 無

(3) 地域分析用に詳細な地理的情報を提供していますか。提供している場合、どのレベルの地理情報が含まれていますか。

地域分析用の地理情報提供の有無：有 無
地理情報のレベル：都道府県、人口 50 万以上の市区

(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。

特定の種類の施設の情報：有 無
項目：「世帯の種類」における「寮・寄宿舎の学生・生徒」、「病院・療養所の入院者」、「社会施設の入所者」、「自衛隊営舎内居住者」及び「矯正施設の入所者」
方法：施設を特定する符号を削除する

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。

世帯の種類、世帯人員、世帯の家族類型、家計の収入の種類（平成 12 年調査）、住居の種類・住宅の所有の関係、住宅の床面積、住宅の建て方、建物全体の階数、世帯が住んでいる階

- (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行っていますか。また、母集団に対する割合はどのようになっていますか。

匿名化措置の方法（項目ごとに記載してください。）

項目：世帯の種類

方法：「一般世帯」及び「施設等の世帯」の2区分を提供する。「施設等の世帯」の内訳は提供しない。

項目：世帯人員

方法：「一般世帯」について、地域区分により、世帯人員が7～9人以上いる世帯を削除。「施設等の世帯」については提供しない。

項目：世帯の家族類型

方法：6区分による提供

項目：家計の収入の種類（平成12年調査）

方法：提供しない。

項目：住居の種類・住宅の所有の関係

方法：住宅の所有の関係について、「公営の借家」及び「都市機構・公社の借家」を統合、「給与住宅」及び「間借り」を統合。「住宅以外に住む一般世帯」の内訳は提供しない。

項目：住宅の床面積

方法：実数ではなく、階級で提供。「200～249 m²」及び「250 m²以上」を統合

項目：住宅の建て方

方法：「長屋建」及び「その他」を統合

項目：建物全体の階数

方法：実数ではなく、階級で提供。高層階の階数区分（「6～10階」、「11～14階」及び「15階以上」）を地域区分によって「6階以上」又は「11階以上」に統合

項目：世帯が住んでいる階

方法：実数ではなく、階級で提供。高層階の階数区分（「3～5階」、「6～10階」、「11～14階」及び「15階以上」）を地域区分によって「3階以上」、「6階以上」又は「11階以上」に統合

(3) 世帯単位の詳細データを提供することに対応して特別な匿名化措置を行っていますか。

特別な措置： <input checked="" type="checkbox"/> 行っている <input type="checkbox"/> 行っていない
項目：父子世帯
方法：削除
項目：年齢差の大きい夫婦のいる世帯
方法：年齢差が25歳以上の夫婦のいる世帯を削除
項目：年齢差の大きい又は小さい親子のいる世帯
方法：年齢差が45歳以上の親と子、年齢差が14歳以下の親と長子又は19歳以下の親と末子のいる世帯を削除
項目：世帯主又は配偶者のいずれか一方若しくは双方が外国人で子供の数が多い世帯
方法：地域区分により、子供の数が3～7人以上いる外国人世帯を削除

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。

年齢、世帯主との続き柄、国籍、5年前の住居の所在地（平成12年調査）、労働力状態、就業時間、従業上の地位、産業(大分類)、職業(大分類)、常住地による従業地・通学地、利用交通手段(平成12年調査)
--

(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行っていますか。また、母集団に対する割合はどのようになっていますか。

匿名化措置の方法（項目ごとに記載してください。）
項目：年齢
方法：5歳階級で提供。85歳以上をトップコーディング
項目：世帯主との続き柄
方法：「他の親族」、「住み込みの雇人」及び「その他」を統合
項目：国籍
方法：「日本人」及び「外国人」の2区分とし、外国籍の内訳は提供しない。
項目：5年前の住居の所在地（平成12年調査）
方法：「他県から」及び「国外から」を統合
項目：労働力状態
方法：「家事などのほか仕事」及び「通学のかたわら仕事」を統合

項目：就業時間

方法：90 時間以上をトップコーディング

項目：従業上の地位

方法：「雇人のある業主」、「雇人のない業主」及び「家庭内職者」を統合

項目：産業(大分類)

方法：「農業」、「林業」及び「漁業」を統合。「鉱業」及び「建設業」を統合。「製造業」及び「電気・ガス・熱供給・水道業」を統合。「複合サービス業」及び「サービス業（他に分類されないもの）」を統合（平成 17 年調査）

項目：職業(大分類)

方法：「保安職業従事者」、「農林漁業作業者」及び「運輸・通信従事者」を統合

項目：常住地による従業地・通学地

方法：「県内他市区町村で従業・通学」及び「他県で従業・通学」を統合

項目：利用交通手段（平成 12 年調査）

方法：「利用交通手段が 1 種類」について、「勤め先・学校のバス」、「ハイヤー・タクシー」、「オートバイ」及び「その他」を統合

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。

誤差の付加： 採用している 採用していない

方法：一部世帯を他の地域の類似世帯と入れ替える。（スワッピング）

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： 行っている 行っていない

方法：母集団の 1% となるように、世帯の種類別に系統抽出を行う。

6 外部の情報

(1) 個人・世帯を特定できる可能性のある外部の情報は存在しますか。

外部情報： 有 無

情報：既存の統計表

方法：母集団一意又は二意であることが判明しているレコードを含む世帯を削除

(2) 母集団情報として利用している情報は何か。

- 国勢調査の調査区名簿
 行政記録から作成した名簿（行政記録の名称： ）
 その他（本調査自体が国勢調査であり、母集団情報として利用している情報は特段ない（調査対象は本邦内に常住している者全て）。）

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。

- 実施している 実施していない
方法：抽出した世帯（1%）を地域区分ごとに世帯単位でランダムに並べ替え

(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。

ない

(3) 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。

5年以上

(4) 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。

ない

国勢調査調査票匿名データの提供項目と匿名化措置一覧
(平成12年)

掲載項目欄の凡例

- ：そのまま提供
●：匿名化措置を講じて提供
－：提供しない

識別情報欄の凡例

- レ：識別情報として考えられる項目

調査項目	提供項目	識別情報	匿名化措置	備考
調査年月	○			
市区町村コード				
都道府県	○			
市区町村	●	レ	・人口50万以上の市区のみ提供	
調査区番号	－	レ		
世帯番号・調査区内連番	－	レ		
世帯員番号	○			
世帯の種類	●	レ	・「一般世帯」及び「施設等の世帯」の2区分を提供 ・「施設等の世帯」の内訳は提供しない	
世帯人員	●	レ	・「一般世帯」については、地域区分により、世帯人員が7～9人以上いる世帯を削除 ・「施設等の世帯」については提供しない	
世帯の家族類型	●	レ	・6区分による提供	
(再掲) 3世代世帯	○			
家計の収入の種類	－			
住居の種類・住宅の所有の関係	●	レ	・「公営の借家」及び「都市機構・公社の借家」を統合 ・「給与住宅」及び「間借り」を統合 ・「住宅以外に住む一般世帯」の内訳は提供しない	
住宅の床面積	●	レ	・実数ではなく、階級で提供 ・「200～249㎡」及び「250㎡以上」を統合	
住宅の建て方	●	レ	・「長屋建」及び「その他」を統合	
共同住宅の階数				
建物全体の階数	●	レ	・実数ではなく、階級で提供 ・高層階の階数区分(「6～10階」、「11～14階」及び「15階以上」)を地域区分によって「6階以上」又は「11階以上」に統合	
世帯が住んでいる階	●	レ	・実数ではなく、階級で提供 ・高層階の階数区分(「3～5階」、「6～10階」、「11～14階」及び「15階以上」)を地域区分によって「3階以上」、「6階以上」又は「11階以上」に統合	
出生の年月				
元号	－	レ		
年	－	レ	・出生年月は年齢に置き換える	
月	－	レ	・5歳階級で提供	
年齢	●	レ	・85歳以上をトップコーディング	
世帯主との続き柄	●	レ	・「他の親族」、「住み込みの雇人」及び「その他」を統合	
男女	○			
配偶関係	○			
国籍	●	レ	・「日本人」及び「外国人」の2区分とし、外国籍の内訳は提供しない	
現在の住居における居住期間	○			
5年前の住居の所在地	●	レ	・「他県から」及び「国外から」を統合	
在学か否かの別	○			
学校の種類・未就学の種類	○			
労働力状態	●	レ	・「家事などのほか仕事」及び「通学のかたわら仕事」を統合	
就業時間				
実数	●	レ	・90時間以上をトップコーディング	
14区分	○			
10区分	○			
(再掲) 40時間以下	○			
(再掲) 48時間以下	○			
従業上の地位	●	レ	・「雇人のある業主」、「雇人のない業主」及び「家庭内職者」を統合	
産業(大分類)	●	レ	・「農業」、「林業」及び「漁業」を統合 ・「鉱業」及び「建設業」を統合 ・「製造業」及び「電気・ガス・熱供給・水道業」を統合	
職業(大分類)	●	レ	・「保安職業従事者」、「農林漁業作業員」及び「運輸・通信従事者」を統合	
常住地による従業地・通学地	●	レ	・「県内他市区町村で従業・通学」及び「他県で従業・通学」を統合	
利用交通手段	●	レ	・「利用交通手段が1種類」について、「勤め先・学校のバス」、「ハイヤー・タクシー」、「オートバイ」及び「その他」を統合	

国勢調査調査票匿名データの提供項目と匿名化措置一覧
(平成17年)

掲載項目欄の凡例

- ：そのまま提供
- ：匿名化措置を講じて提供
- －：提供しない

識別情報欄の凡例

- レ：識別情報として考えられる項目

調査項目	提供項目	識別情報	匿名化措置	備考
調査年月	○			
市区町村コード				
都道府県	○			
市区町村	●	レ	・人口50万以上の市区のみ提供	
調査区番号	－	レ		
世帯番号・調査区内連番	－	レ		
世帯員番号	○			
世帯の種類	●	レ	・「一般世帯」及び「施設等の世帯」の2区分を提供 ・「施設等の世帯」の内訳は提供しない	
世帯人員	●	レ	・「一般世帯」については、地域区分により、世帯人員が7～9人以上いる世帯を削除 ・「施設等の世帯」については提供しない	
世帯の家族類型	●	レ	・6区分による提供	
(再掲) 3世代世帯	○			
住居の種類・住宅の所有の関係	●	レ	・「公営の借家」及び「都市機構・公社の借家」を統合 ・「給与住宅」及び「間借り」を統合 ・「住宅以外に住む一般世帯」の内訳は提供しない	
住宅の床面積	●	レ	・実数ではなく、階級で提供 ・「200～249㎡」及び「250㎡以上」を統合	
住宅の建て方	●	レ	・「長屋建」及び「その他」を統合	
共同住宅の階数				
建物全体の階数	●	レ	・実数ではなく、階級で提供 ・高層階の階数区分(「6～10階」、「11～14階」及び「15階以上」)を地域区分によって「6階以上」又は「11階以上」に統合	
世帯が住んでいる階	●	レ	・実数ではなく、階級で提供 ・高層階の階数区分(「3～5階」、「6～10階」、「11～14階」及び「15階以上」)を地域区分によって「3階以上」、「6階以上」又は「11階以上」に統合	
出生の年月				
元号	－	レ	・出生年月は年齢に置き換える	
年	－	レ	・5歳階級で提供	
月	－	レ	・85歳以上をトップコーディング	
年齢	●	レ		
世帯主との続き柄	●	レ	・「他の親族」、「住み込みの雇人」及び「その他」を統合	
男女	○			
配偶関係	○			
国籍	●	レ	・「日本人」及び「外国人」の2区分とし、外国籍の内訳は提供しない	
労働力状態	●	レ	・「家事などのほか仕事」及び「通学のかたわら仕事」を統合	
就業時間				
実数	●	レ	・90時間以上をトップコーディング	
14区分	○			
10区分	○			
(再掲) 40時間以下	○			
(再掲) 42時間以下	○			
(再掲) 48時間以下	○			
従業上の地位	●	レ	・「雇人のある業主」、「雇人のない業主」及び「家庭内職者」を統合	
産業(大分類)	●	レ	・「農業」、「林業」及び「漁業」を統合 ・「鉱業」及び「建設業」を統合 ・「製造業」及び「電気・ガス・熱供給・水道業」を統合 ・「複合サービス業」及び「サービス業(他に分類されないもの)」を統合	
職業(大分類)	●	レ	・「保安職業従事者」、「農林漁業作業員」及び「運輸・通信従事者」を統合	
常住地による従業地・通学地	●	レ	・「県内他市区町村で従業・通学」及び「他県で従業・通学」を統合	

平成12年国勢調査の概要

調査の目的

国勢調査は、我が国の人口、世帯、産業構造等の実態を明らかにし、国及び地方公共団体の各種行政施策の基礎資料を得ることを目的として行われる国の最も基本的な統計調査である。調査は大正9年以来ほぼ5年ごとに行われており、平成12年国勢調査はその17回目に当たる。

調査の時期

平成12年国勢調査は、平成12年10月1日午前零時(以下「調査時」という。)現在によって行われた。

調査の法的根拠

平成12年国勢調査は、統計法(昭和22年法律第18号)第4条第2項の規定並びに次の政令及び総理府令に基づいて行われた。

国勢調査令(昭和55年政令第98号)

国勢調査施行規則(昭和55年総理府令第21号)

国勢調査の調査区の設定の基準等に関する総理府令(昭和59年総理府令第24号)

調査の地域

平成12年国勢調査は、我が国の地域のうち、国勢調査施行規則第1条に規定する次の島を除く地域において行われた。

- (1) 歯舞群島、色丹島、国後島及び択捉島
- (2) 島根県隠岐郡五箇村にある竹島

調査の対象

平成12年国勢調査は、調査時において、本邦内に常住している者について行った。ここで「常住している者」とは、当該住居に3か月以上にわたって住んでいるか、又は住むことになっている者をいい、3か月以上にわたって住んでいる住居又は住むことになっている住居のない者は、調査時現在いた場所に「常住している者」とみなした。

ただし、次の者については、それぞれ次に述べる場所に「常住している者」とみなしてその場所で調査した。

- 1 学校教育法(昭和22年法律第26号)第1条に規定する学校、第82条の2に規定する専修学校又は第83条第1項に規定する各種学校に在学している者で、通学のために寄宿舎、下宿その他これらに類する宿泊施設に宿泊している者は、その宿泊している施設
- 2 病院又は療養所に引き続き3か月以上入院し、又は入所している者はその入院先、それ以外の者は3か月以上入院の見込みの有無にかかわらず自宅
- 3 船舶(自衛隊の使用する船舶を除く。)に乗り組んでいる者で陸上に生活の本拠を有する者はその生活の本拠である住所、陸上に生活の本拠のない者はその船舶

なお、後者の場合は、日本の船舶のみを調査の対象とし、調査時に本邦の港に停泊している船

船のほか、調査時前に本邦の港を出港し、途中外国の港に寄港せず調査時後5日以内に本邦の港に入港した船舶について調査した。

- 4 自衛隊の営舎内又は自衛隊の使用する船舶内の居住者は、その営舎又は当該船舶が籍を置く地方総監部(基地隊に配属されている船舶については、その基地隊本部)の所在する場所
- 5 刑務所、少年刑務所又は拘置所に収容されている者のうち、死刑が確定した者及び受刑者並びに少年院又は婦人補導院の在院者は、その刑務所、少年刑務所、拘置所、少年院又は婦人補導院

本邦内に常住している者は、外国人を含めてすべて調査の対象としたが、次の者は調査の対象から除外した。

- (1) 外国政府の外交使節団・領事機関等の構成員(随員を含む。)及びその家族
- (2) 外国軍隊の軍人・軍属及びその家族

調査の事項

平成12年国勢調査では、男女の別、出生の年月など世帯員に関する事項を16項目、世帯の種類、世帯員の数など世帯に関する事項を6項目、合計22項目について調査した。

(世帯員に関する事項)

- | | |
|------------------|------------------------|
| (1) 氏名 | (9) 在学、卒業等教育の状況 |
| (2) 男女の別 | (10) 就業状態 |
| (3) 出生の年月 | (11) 就業時間 |
| (4) 世帯主との続柄 | (12) 所属の事業所の名称及び事業の種類 |
| (5) 配偶の関係 | (13) 仕事の種類 |
| (6) 国籍 | (14) 従業上の地位 |
| (7) 現在住居における居住期間 | (15) 従業地又は通学地 |
| (8) 5年前の住居の所在地 | (16) 従業地又は通学地までの利用交通手段 |

(世帯に関する事項)

- | | |
|--------------|------------|
| (1) 世帯の種類 | (4) 住居の種類 |
| (2) 世帯員の数 | (5) 住宅の床面積 |
| (3) 家計の収入の種類 | (6) 住宅の建て方 |

調査の方法

平成12年国勢調査は、総務庁(統計局・統計センター)－都道府県－市町村－国勢調査指導員－国勢調査員の流れにより行った。

調査は、総務庁長官により任命された約83万人の国勢調査員が調査票を世帯ごとに配布し、収集する方法により行った。また、調査票への記入は、原則として世帯が行った。

集計及び結果の公表

集計は、総務省(平成13年1月5日以前は総務庁)統計センターが行い、結果の公表は、総務省統計局が結果原表を閲覧に供する方法等により行う。また、主要な結果を収録した報告書を結果の公表から約1か月後に刊行する。

平成17年国勢調査の概要

調査の目的

国勢調査は、我が国の人口、世帯、産業構造等の実態を明らかにし、国及び地方公共団体の各種行政施策の基礎資料を得ることを目的として行われる国の最も基本的な統計調査である。調査は大正9年以来ほぼ5年ごとに行われており、平成17年国勢調査はその18回目に当たる。

調査の時期

平成17年国勢調査は、平成17年10月1日午前零時（以下「調査時」という。）現在によって行われた。

調査の法的根拠

平成17年国勢調査は、統計法（昭和22年法律第18号）第4条第2項の規定並びに次の政令及び総理府令に基づいて行われた。

国勢調査令（昭和55年政令第98号）

国勢調査施行規則（昭和55年総理府令第21号）

国勢調査の調査区の設定の基準等に関する総理府令（昭和59年総理府令第24号）

調査の地域

平成17年国勢調査は、我が国の地域のうち、国勢調査施行規則第1条に規定する次の島を除く地域において行われた。

- (1) 歯舞群島、色丹島、国後島及び択捉島
- (2) 島根県隠岐郡隠岐の島町にある竹島

調査の対象

平成17年国勢調査は、調査時において、本邦内に常住している者について行った。ここで「常住している者」とは、当該住居に3か月以上にわたって住んでいるか、又は住むことになっている者をいい、3か月以上にわたって住んでいる住居又は住むことになっている住居のない者は、調査時現在居た場所に「常住している者」とみなした。

ただし、次の者については、それぞれ次に述べる場所に「常住している者」とみなしてその場所で調査した。

- 1 学校教育法（昭和22年法律第26号）第1条に規定する学校、第82条の2に規定する専修学校又は第83条第1項に規定する各種学校に在学している者で、通学のために寄宿舍、下宿その他これらに類する宿泊施設に宿泊している者は、その宿泊している施設
- 2 病院又は療養所に引き続き3か月以上入院し、又は入所している者はその病院又は療養所、それ以外の者は3か月以上入院の見込みの有無にかかわらず自宅
- 3 船舶（自衛隊の使用する船舶を除く。）に乗り組んでいる者で陸上に生活の本拠を有する

者はその住所、陸上に生活の本拠の無い者はその船舶

なお、後者の場合は、日本の船舶のみを調査の対象とし、調査時に本邦の港に停泊している船舶のほか、調査時前に本邦の港を出港し、途中外国の港に寄港せず調査時後5日以内に本邦の港に入港した船舶について調査した。

4 自衛隊の営舎内又は自衛隊の使用する船舶内の居住者は、その営舎又は当該船舶が籍を置く地方総監部（基地隊に配属されている船舶については、その基地隊本部）の所在する場所

5 刑務所、少年刑務所又は拘置所に収容されている者のうち、死刑が確定した者及び受刑者並びに少年院又は婦人補導院の在院者は、その刑務所、少年刑務所、拘置所、少年院又は婦人補導院

本邦内に常住している者は、外国人を含めてすべて調査の対象としたが、次の者は調査の対象から除外した。

- (1) 外国政府の外交使節団・領事機関の構成員（随員を含む。）及びその家族
- (2) 外国軍隊の軍人・軍属及びその家族

調査の事項

平成17年国勢調査では、男女の別、出生の年月など世帯員に関する事項を12項目、世帯の種類、世帯員の数など世帯に関する事項を5項目、合計17項目について調査した。

（世帯員に関する事項）

- | | |
|--------------|----------------------|
| (1) 氏名 | (7) 就業状態 |
| (2) 男女の別 | (8) 就業時間 |
| (3) 出生の年月 | (9) 所属の事業所の名称及び事業の種類 |
| (4) 世帯主との続き柄 | (10) 仕事の種類 |
| (5) 配偶の関係 | (11) 従業上の地位 |
| (6) 国籍 | (12) 従業地又は通学地 |

（世帯に関する事項）

- | | |
|-----------|------------|
| (1) 世帯の種類 | (4) 住宅の床面積 |
| (2) 世帯員の数 | (5) 住宅の建て方 |
| (3) 住居の種類 | |

調査の方法

平成17年国勢調査は、総務省（統計局）－都道府県－市町村－国勢調査指導員－国勢調査員の流れにより行った。

調査は、総務大臣により任命された約85万人の国勢調査員が調査票を世帯ごとに配布し、収集する方法により行った。また、調査票への記入は、原則として世帯が行った。

集計及び結果の公表

集計は、独立行政法人統計センターが行い、結果の公表は、総務省統計局が結果原表を閲覧に供する方法等により行う。また、主要な結果を収録した報告書を結果の公表から約1か月後に刊行する。

(1) 情報の削除

- ア： 識別情報の削除等** 直接的な識別情報は、レコードから全面的に削除する。
また、レコードの配列順が意味をなさないように、無作為に並べ替えを行う。
- イ： 裾切りによるレコード削除** 発生頻度の低いレコード及び特徴的な値があるレコードを含む世帯を削除する。
- ウ： 母集団一意等の削除** 既存の統計表により母集団一意又は二意であることが判明しているレコードを含む世帯を削除する。

(2) データの抽出， 識別情報の階級区分統合， 地域情報の入替

- ア： サンプルング** 全世帯数(母集団)から1%の割合で抽出したデータを用いる。
- イ： トップコーディング** 極端に大きな値は、上限値を設けて頭打ちにする。
- ウ： リコーディング** 分類区分の程度は、詳細なものではなく、粗いものとする。
- エ： スワッピング** 一部世帯を他の地域の類似世帯と入れ替える。

