

平成 19 年国民生活基礎調査の匿名データ 仕様及び利用上の注意  
(データ A 及び B 共通)

平成 19 年国民生活基礎調査の匿名データの利用に際しては、以下の事項を十分理解した上で集計、分析等を行ってください。

## 1 全般的注意事項

### (1) 本調査に関する理解

利用に際しては、報告書、当省ホームページ及び「政府統計の総合窓口」e-stat 等により、本調査の説明、調査票様式、標本設計、調査方法等を確認し、十分に理解してください。これらの内容に関する理解が十分でないと、匿名データの集計、分析に誤りが生じる可能性があります。

#### <関係リンク>

- 厚生労働省 厚生労働統計一覧トップページ

<http://www.mhlw.go.jp/toukei/itiran/>

- ・ 「平成 19 年国民生活基礎調査の概況」

<http://www.mhlw.go.jp/toukei/list/20-19-1.html>

… 概要と主な統計表のみを掲載しているため、詳細な内容は報告書又は「政府統計の総合窓口」e-stat をご覧ください。

- ・ 調査票様式

<http://www.mhlw.go.jp/toukei/chousahyo/index.html>

- 「政府統計の総合窓口」e-stat

<http://www.e-stat.go.jp/SG1/estat/eStatTopPortal.do>

… 詳細な統計表をデータベース化しています。

※ 本書類に記載しているリンクは、アドレスが変更される場合があります。

### (2) 調査時点等

#### 1) 世帯票及び健康票

調査日はいずれも平成 19 年 6 月 7 日で、次の項目を除き調査日現在の状況です。

世帯票のうち家計支出総額、育児にかかった費用、仕送りの費用、仕事の状況、1 週間の就業日数等、健康票のうち医療費支払総額は、平成 19 年 5 月中の状況です。

#### 2) 所得票及び貯蓄票

調査日はいずれも平成 19 年 7 月 12 日です。

所得票の項目は、平成 18 年 1 月 1 日から 12 月 31 日までの 1 年間の状況、貯蓄票の項目は、平成 19 年 6 月末日現在の状況です。

### (3) 公表結果との差異

匿名データは、被調査者が特定されないよう、リサンプリングやトップコーディング等の秘匿措置を施している（「5 秘匿措置の適用」参照）ため、匿名データによる集計結果は、当省で公表している結果と完全には一致しません。

### (4) その他

- 1) 調査票の記入誤りなどはチェック・修正済みですが、被調査者の回答に基づくものであるため、項目間に論理的な整合性がとれていない場合があります。
- 2) 後述のとおり、被調査者が特定される恐れのある世帯のレコードは削除していますが、外れ値を除外する処理は行っていないため、多変量解析等を行う際には、外れ値を考慮してください。
- 3) 提供資料（符号表及びデータレイアウト等）の記載内容をよく理解した上で匿名データを利用してください。

## 2 データの形式等

### (1) データ形式

データは CSV 形式です。

なお、データは、全てのレコードのレコード長及び項目のカラム数が一致するよう、固定長テキスト形式のレコードをカンマで区切っているため、固定長テキスト形式としても扱えます。

文字コードは Shift\_JIS ですが、漢字、カタカナ等は使用していません。

改行コードは CRLF です。

### (2) レコード構成

1 世帯員 1 レコードの構成としています。

レコードは、世帯番号、世帯員番号の順に並べているため、世帯への統合が可能です。また、同一世帯番号の世帯員には、すべて同一の世帯共通項目を複製しています。

（したがって、世帯で集計、分析する場合には、原則として世帯員番号「01」の世帯主のレコードに着目して処理します。詳細は 4（2）参照）

表 1 レコード構成の概要

データ名	データ A	データ B
データファイル名	H19KOKUMIN_A.CSV	H19KOKUMIN_B.CSV
レコード件数 (1レコード1世帯員)	98,789	14,293
(参考) 世帯数換算	37,676 世帯	5,282 世帯

データ容量（メガバイト）	66	9
レコード長（バイト）	688	688
世帯票及び健康票の項目	提供項目は同一 (ただし、提供レコードはそれぞれ異なる)	
所得票及び貯蓄票の項目	なし	提供
ウェイト	提供	なし

※項目は、秘匿性確保の観点から、一部に提供しないもの、秘匿措置を講じた上で提供する場合があります。詳細は符号表をご覧ください。

### 3 符号表について

#### (1) 符号表の記法

符号表は「政府統計個票データレイアウト標準記法」（平成 18 年 3 月 31 日各府省情報化統括責任者（CIO）連絡会議決定。以下、「標準記法」という。）に概ね準拠しているため、標準記法を確認の上、利用してください。

- 総務省統計局 標準記法リンク

<http://www.mhlw.go.jp/toukei/itaku/dl/layout.pdf>

#### (2) 不詳項目

項目の内容が不詳の場合、標準記法とは異なり、「9」、「99」、「999」等を用いて表記しているため、必ず符号表で確認してください。

#### (3) 桁数

データの符号については、2桁以上の項目は0又はブランク（符号表では△）の補足により桁数を揃えてあります。

#### (4) 分類項目

調査項目から作成し、統計表作成に用いている主な分類項目（「世帯類型」等）を付与しています。定義は報告書等を参照してください。

#### (5) データA及びBのデータレイアウト

データA及びBによる提供項目は一部異なりますが、双方のデータレイアウトは共通としているため、提供しない項目のカラムは、ブランク（符号表上は「△」）となっています。

### 4 ウェイト

#### (1) ウェイトとは

ウェイトとは、標本調査で母集団の値を推定するために、各レコード（標本）に乘じ

る係数（個別データの重み）で、抽出率等を加味した値です。

## (2) 匿名データのウェイト

### 1) データ A

全国の推計世帯数の算出が可能となるよう、一律のウェイトを付与しています。

ウェイトは、1,000 倍した整数値で格納しているため、1000 で除して利用してください。

推計値の算出方法は次のとおりです。

$$\bigcirc \text{ 推計世帯数} = \text{世帯員番号「01」のレコード件数} \times \text{ウェイト}$$

$$\bigcirc \text{ 推計人数} = \text{レコード件数} \times \text{ウェイト}$$

(参考：提供しているウェイトの算出方法)

$$\begin{aligned} \text{ウェイト} &= \text{平成 19 年 6 月 1 日現在 日本人推計人口 (総務省統計局)} \\ &\quad / \text{ データ A の全レコード件数} \\ &= 126,082,831 / 98,789 \times 1,000 \\ &\approx 1,276,284 \end{aligned}$$

### 2) データ B

推計世帯数の算出を前提としていないため、ウェイトは付与していません。

(所得票及び貯蓄票の公表結果においては、推計世帯数は算出していません。)

## 5 秘匿措置の適用

匿名データは、統計法第 2 条第 12 項により、「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」することが定められていることを踏まえ、本匿名データでは、以下の秘匿措置を講じているため、留意して利用してください。

なお、これらの秘匿措置は、統計法第 35 条第 2 項に基づいて、内閣府統計委員会（匿名データ部会）における有識者の審議を経て、答申、了承されたものです。

### (1) 地域

地域は「全国」のみとしています。

### (2) リサンプリング

リサンプリングは、国勢調査区（又は単位区）及び世帯の二段階で行い、各レコードの重みが一律になるよう調整しています。

リサンプリング後のサンプルサイズは、概ね中間年と同程度です。

(本調査は毎年実施していますが、3年に1度大規模調査を実施し、大規模調査年以外の調査年を中間年といたします。平成 19 年は大規模調査年に当たります。)

### (3) 年齢階級

年齢は原則として 5 歳階級とし、90 歳以上は同一の階級としています。(トップコー

ディング)

15歳未満の世帯員の年齢は、健康票の記入対象年齢区分を踏まえて階級化(0～5歳、6～11歳、12～14歳)しています。(リコーディング)

**(4) レコード削除**

世帯員が8人以上の世帯、父子世帯、手助けや見守りを要する者が2人以上いる世帯、要介護認定を受けている者が2人以上いる世帯、年齢階級差の大きい夫婦がいる世帯、年齢階級差の大きい又は小さい親子がいる世帯、同一年齢階級に4人以上がいる世帯は、リサンプリング前に世帯単位でレコードを削除しています。

**(5) レコードの並び**

レコードは世帯単位で無作為に並び替えています。

**(6) リコーディング**

一部の項目の選択肢については、上位区分への統合、他の選択肢との統合、「その他」への統合を行っています(詳細は符号表参照。)

**(7) トップコーディング、ボトムコーディング**

一部の項目については、上限値、下限値を設け、上限値以上、下限値以下は、当該上限値(下限値)に置き換えて提供します(符号表参照)。

したがって、トップコーディング又はボトムコーディングしている階級を含む分析、平均値の算出、分布の観察、多変量解析等の際には、これらの影響に留意してください。

**(8) 所得項目の内訳**

世帯の所得の種類別内訳項目及び世帯員別の所得額については、秘匿性を確保するため提供していません。

**6 その他**

**(1) 統計法等の遵守**

統計法、提供依頼申出書及び約款による規定内容を遵守し、特に適正管理には十分留意の上、利用してください。

**(2) 管理状況報告書の提出**

匿名データの利用期間が1年間を超える場合は、定期的に管理状況報告書を提出してください。

**(3) 変更手続**

提供依頼申出書に記載した内容に変更が生じる場合は、事前にご相談ください。

**(4) 成果の公表**

成果の公表に際しては、必ず以下の内容を明示してください。

- 1) 「統計法第36条の規定に基づいて厚生労働省から提供を受けた匿名データを利用したこと(「マイクロデータ」や「個票」ではなく、「匿名データ」と記載してください。)
- 2) 匿名データを利用して得られた結果は、匿名データを基に利用者が独自に作成、加工した統計等であり、厚生労働省が作成、公表しているものとは異なること

(5) 匿名データ利用終了後の措置

利用期間終了日までに、次の1)～3)を窓口までご提出ください。

- 1) 提供を受けた匿名データ
- 2) データ消去等報告書
- 3) 利用実績報告書

(6) 担当窓口

ご質問等は、以下の担当窓口へご連絡ください。

厚生労働省 政策統括官（統計・情報政策、政策評価担当）付

参事官付 審査解析室

<メールでのお問い合わせ>

E-mail : [nijitekiryoyou@mhlw.go.jp](mailto:nijitekiryoyou@mhlw.go.jp)

<お電話でのお問い合わせ>

平日 10:00～17:00（12:00～13:00 を除く）

代表：03-5253-1111 内線 7391