

第6回共通課題検討ワーキンググループ

調査横断的且つ一般的な匿名化のルール等について (案)

令和4年11月22日 統計研究研修所

1. 匿名データの作成について（現行）

調査ごとに定められている匿名化処理基準に基づき作成

➤ 調査票情報の匿名化処理の基準

各調査、各年次の匿名データの作成に合わせて、統計委員会において審議されて「匿名データの作成に係る匿名化処理基準」が定められている

➤ 匿名データの作成の検討に当たっての流れ

● 匿名データの作成方針等を作成

• 匿名データの作成方針

匿名データ作成計画、匿名化処理基準等を踏まえ、基本的な考え方、作成する匿名データの構成、適用する匿名化処理等を記載

• 匿名データの審査表

識別情報、調査項目などその実施する具体的な匿名化処理の方法等を記載

• 匿名化の基準に適合しているかどうかを確認するための資料

度数分布表等を作成

● 有識者による審査

匿名データ有識者会議（統計研究研修所）において、上記匿名データの作成方針等を審査

● 統計委員会において審議、承認

2. 調査横断的且つ一般的な匿名化のルール

各調査で共通な匿名化処理とすることが可能であるものをルールとして整備

※ このルールとして整理できない事項については、従前どおり調査の特性に応じて各調査において検討する

本日、議論いただく事項

(1) リサンプリング (サンプリング)

- 現在、リサンプリングの割合が80%且つ報告者の数が多い調査を対象とする
- 構成割合、基本的な統計量等に問題がなければ、割合を下げる
(なお、各調査の匿名データの作成検討時に匿名化処理の見直しが生じた場合は、割合を再検討する)

(2) 削除する情報 (世帯等のレコード単位での削除)

- 世帯人員数の多い世帯、同一年齢の子供の多い世帯等の発生頻度の少ないものを削除する

(3) 外観識別性

- 調査間共通の定義として、「客観的且つ視覚的に識別可能なもの」とする
⇒ 「住宅関係 (現住居に関するもの)」の調査項目

(4) 調査項目の匿名化

①提供しない調査項目

- 出生の年月(注：年齢に換算して提供)、特定の災害に関する調査項目等は提供しない

②そのまま提供する調査項目

- 調査の特性がある調査項目、意識を問う調査項目等は、原則そのまま提供する

▶ 上記①、②に該当しない調査項目については匿名化の検討を行う

3. 匿名化処理に関する情報提供の在り方

➤ 現状

匿名データの作成方法、処理される値等を「匿名データの審査表」に具体的に記載し、その内容を匿名データ有識者会議及び統計委員会のHP等で公開している

➤ 今後の情報提供の在り方

利用に当たって必要な匿名化処理の情報と秘匿性の確保の双方を考慮しつつ、情報提供を行う

- 現行の審査表に記載されている調査項目ごとの具体的な処理方法、処理される値等については非公開
- 匿名データの利用を検討するに当たって必要な情報を提供（調査項目ごとの匿名化処理の情報）
- 作成方法の概要等を記載した資料を新たに作成（現行の「匿名データの作成方針」及び「匿名データの審査表」を統合）
 - ・ しきい値等の具体的な数値は掲載しない
 - ・ 調査横断的且つ一般的な匿名化のルールと併せて、公開する資料とする

(参考) 調査横断的且つ一般的な匿名化のルールを進め方

※ 調査横断的且つ一般的な匿名化のルールについて、以下の内容を2回に分けて共通課題検討WGで検討を進める

- 第6回 (令和4年11月22日)
 - リサンプリング (サンプリング)
 - 削除する情報 (世帯等のレコード単位での削除)
 - 外観識別性
 - 調査項目の匿名化
 - 提供しない調査項目
 - そのまま提供する調査項目
 - 匿名化処理に関する情報提供の在り方

- 第7回 (令和5年1月予定)
 - 調査項目の匿名化 (匿名化の検討を行うもの)
 - しきい値による処理 (0.5%基準)
 - 連続変数扱いをする調査項目の指定
 - 分類区分の統合の方針及び留意点