

保存期間：5年  
(令和10事務年度末)  
令和5年11月28日

## 第7回 国税庁保有行政記録情報の整備に関する技術検証WG

### 議事要旨

日時：令和5年11月28日（火）10：00～12：05

場所：国税庁会議室

出席者：伊藤伸介座長、菅幹雄委員、星野申明委員、南和宏委員、国税庁企画課職員

国税庁企画課から、配付資料に基づき、検討内容について説明。その後、以下のとおり委員から意見があった。

#### ○ サンプルングについて

- ・ 1%は妥当な水準ではないか。1%のサンプルサイズ(約23万レコード)があれば、実証分析を十分に行うことはできるのではないか。仮により多くのデータが必要な場合には、税務大学校との共同研究の枠組みで対応が可能であろう。
- ・ 安全性の面から言えば、サンプルング率は小さいほど良い。
- ・ サンプルング方法は、安全性の観点からは非復元抽出より復元抽出の方が好ましい。
- ・ サンプルング率や方法の検討に当たっては、利用者ニーズも考慮する必要がある。

#### ○ 識別性の高い情報

- ・ 地域情報は外観識別性が高いため、加工方法は特に注意する必要がある。詳細な地方区分での提供は安全性の観点から困難だろう。一方、地域によっては、細かな区分による提供が可能の場合も考えられるが、他の情報の提供方法との組合せを踏まえたより詳細な検討が必要である。
- ・ 業種情報は、分析する上で重要な情報であると考えられるため、分類区分を統合する場合には慎重な検討が必要であろう。
- ・ 年齢情報を5歳階級で提供する場合、利用の観点からも特に違和感はない。
- ・ 性別情報について、令和2年分以降はデータを保持していなくとも、既存のデータは提供した方が良い。
- ・ 地域情報や年齢、業種等の他の情報と組み合わせた場合のレコード数に応じて、適切な匿名加工を検討する必要がある。

○ その他の情報

- ・ 定量情報を1円単位で提供することはリスクが高まるため、ラウンディングやグルーピングは行うべきであろう。
- ・ 定量情報に合計項目と内訳項目がある場合、全ての内訳項目の情報を提供する必要があるかどうかは、安全性の観点から検討が必要ではないか。
- ・ 利用者にとっては、匿名データに攪乱的手法が使われることは好ましくない可能性があるため、まずは攪乱的手法を使わない加工方法を検討してはどうか。
- ・ 匿名データの各項目の標準誤差率や、トップコーディングを適用した部分の平均値や分散等の情報を提供してはどうか。

以 上