

## 第 10 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 7 年 11 月 4 日（火） 13:00～16:00

場 所：国税庁会議室

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

冒頭、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した後、意見交換を実施。委員からは、税務データの匿名加工について以下の方向性で作業を進めることが適当との意見があつた。

#### ○ 地域情報の提供

- ・ 住所の粒度は、47都道府県単位とする
- ・ 特定リスクの高いレコードは、平均値に置き換えて提供することとし、当該平均値は、局所平均を取って提供する
- ・ 平均値に置き換えたレコードを判別できるようにフラグをつける
- ・ 抽出率は 1% とする
- ・ 準識別子（年齢・住所・業種）のクロスごとの平均値・標準偏差を統計量として別途提供する
- ・ 上記を前提とした匿名データは、現行の匿名データから置き換えて提供する

#### ○ データのパネル化

- ・ まずは、安全性が高いと考えられる以下の方法により提供を開始し、提供後のニーズを踏まえ、必要に応じて有用性の向上を検討する
- ・ パネル化の情報範囲については、準識別子を年齢に絞った上で提供する
- ・ パネル化の年数は、連続した 2 年間とする
- ・ 連続した 2 年間のデータをパネルデータと捉え、1% 標本抽出する
- ・ 年齢情報のグルーピング、定量情報のトップコーディングやラウンディング、内訳削除の処理を継続するが、コーディングについては、安全性の観点から 0.5% から 1% に基準を引き上げる
- ・ 都道府県レベルで提供する各年データとパネルデータは、分けて提供することとし、データのリンクエージは認めないこととする

## ○ その他

- 同一のレコードが2年間のパネルデータに重複してサンプリングされても安全性には問題はない。サンプリングする際は、データの分布に歪みが生じないよう留意すべき
- 統計的マッチング (Statistical Matching) は、個人を特定するためのリンクエージには当たらないため、認めてよいのではないか
- 利用申請に必要な書類の様式について、作成手続きを簡略化するため、リサーチマップを利用する場合には、記載内容が正確であるかどうかを確認した方がよい
- 利用目的の「税財政政策に関する」という制限については、将来的に緩和する方向で検討してはどうか

以 上