

保存期間：5年
(令和10事務年度末)
令和6年4月9日

第8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6年4月9日（火）13：30～15：00

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

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。

- 識別性の高い情報（定性情報）
 - ・ 識別性の高い情報を全て提供する場合は特定リスクが高まるため、いずれかの情報は削除する必要がある、その場合、性別を削除する方法が最も合理的という事務局案に賛成である。
- その他の情報（定量情報）
 - ・ 出現率の低い項目については、レコードごと削除されても仕方ないだろう。一方、仮に削除した場合、分布にゆがみが生じる可能性もあるため、事後検証が必要である。
 - ・ 収入金額等をグルーピングすると分析の精度が落ちてしまうため、利用者は嫌がるのではないか。
 - ・ 内訳の金額をグルーピングしても、合計額から推計できるという点は、そこまで慎重になる必要は無いのではないか。その数値や分布から個人が特定されるか否かが問題で、ケースバイケースで考えるしかない。
 - ・ 内訳の金額から合計額が推計でき、合計額のトップコーディングが機能しないという点は、内訳の金額を%表示にすれば問題は解決できるが、一方で、実際の税務データは%では保持されていないため、利用者がどのようなイメージを持つか気になる。その点、トップコーデコーディングをした合計額のしきい値から各内訳の金額を割り返すという方法が一案として考えられるが、有用性の観点から、分布にどの程度影響が出るかは検証が必要である。
- その他（提供対象年分、サンプリング方法、サンプルデータ）
 - ・ 古い年分から提供するという事務局案に賛成である。
 - ・ サンプリング方法は、最終的には実用性の観点から決めてはどうか。

- ・ サンプルデータは、まず利用者がデータ構造を理解するためのメタデータとして提供するものの、教育目的で使用することも可能ではないか。ただし、教育用途で有用性のあるデータを作るためには、それなりの技術力が必要であるため、事前に方針を決めておいた方が良い。

以 上