

委託先におけるプログラムミスによる 統計数値の誤り発生について

(追加調査の結果)

令和元年7月19日
統計委員会担当室

プログラムミスによる結果数値訂正の事案(原因等の詳細は後述)

これまでの書面調査等で報告のあった結果数値訂正案件について、詳細を追加調査により確認した結果(調査事項は末尾の参考資料を参照)、以下のとおり、受託事業者のプログラムミスと考えられる事案が確認された(計13調査14事案)。

なお、プログラム外の人的ミスであることが明らかな事案(使用名簿の誤り、単純な入力ミス など)、結果数値に直接影響しない事案(集計表見出しの記載誤り など)、統計作成部局外の業務に係る事案(集計に用いる行政記録作成時の誤り)はカウントから除外した。

集計データの取り込み(1)

○電子レセプトデータを集計システムに取り込む際の設定誤りにより、一部データが正しく取り込まれなかった(1事案)

①医療扶助実態調査(厚生労働省) **仕**

(凡例)

仕 当該作業内容が仕様書等上、不明確(p.3,4)
※上記以外の事案では明確

テ テストデータを用いた検証を実施せず(p.8)

テ 何らかのテストデータを用いた検証を実施していた事案(p.8)

集計データの作成

(集計データ自体の誤り)(9)

○産業格付けの分類番号を誤って変更した(1事案)

②情報通信業基本調査(総務省、経済産業省の共管) **テ**

○集計対象の範囲指定に誤りがあった(正「<」、誤「≤」等)(5事案)

③国勢調査(総務省) **テ**

④通信利用動向調査(総務省) **テ**

⑤健康保険・船員保険被保険者実態調査(厚生労働省) **仕**

⑥障害福祉サービス等従事者処遇状況等調査(厚生労働省) **仕**

⑦中小企業実態基本調査(中小企業庁) **仕**

○別々に集計すべき項目のうち一つが二重に計上された(1事案)

⑧大学等におけるフルタイム換算データに関する調査(文部科学省) **テ**

○記載されるべき数値の設定に誤りがあった(0と空欄の区別 等)(1事案)

⑨学術情報基盤実態調査(文部科学省) **仕**

○プログラムにおいて、処理する変数を書き間違えた(1事案)

⑩訪日外国人消費動向調査(観光庁) **テ**

公表資料作成

(データ自体でなく表示箇所の誤り)(4)

○結果表(公表資料)を作成する際にデータの参照先設定を誤ったため、表内の結果数値が正しく表示されなかった(4事案)

⑪学術情報基盤実態調査(文部科学省) **テ**

⑫建設工事統計調査(建設工事施工統計調査)(国土交通省) **テ**

⑬国際航空旅客動態調査(国土交通省) **テ**

⑭全国貨物純流動調査(国土交通省) **テ**

(注)学術情報基盤実態調査は1調査で2事案をカウント

具体例

○結果表(公表資料)を作成する際に、データの参照先設定を誤ったため、表内の結果数値が正しく表示されなかった(4事案)

【事例】⑬国際航空旅客動態調査(国土交通省)

発見した契機	事案の概要	仕様書等との関係	履行確認の実施状況
公表後、ユーザーからの結果内容に関する問い合わせを受けて、再度公表データの見直しを行ったところ、誤字や数値の誤りが判明。	プログラムの一部に誤りがあったため、本来参照されるべき数値が正しく参照されていないことが等から、該当の3表を修正。	仕様書等上で作成すべき結果表は明示されており、委託先も作成すべき結果表の理解はあったが、プログラマーの作業ミスで結果表に対応するデータの参照先設定が違ってしまった。	・テストデータによる検証は行っていない。 ・一次納品物(仮納品)による確認作業は、空港毎の合計値に大幅な変化がないかという観点で確認を行うため、納品物(PDF)の報告としては確認できない部分(プログラムによって参照される数値)の誤りまで発見できなかった。

異なる列に同じ数値が記載されているケース

本来集計されるべきセルの値が記載されていないケース

23 最終アクセス交通手段
■上段：旅客数 下段：構成比

	成	田	西	京
空港発行動バス	97,570	72,788	5,623	31.5
高速度バス(空港発行動バス以外)	-	-	-	-
貸切りバス	33,981	28,263	14,329	11.0
路線バス	391	1,142	84	0.1
新幹線	-	-	-	-
優等列車	61,123	46,717	19,877	19.7
鉄道・地下鉄	78,584	95,030	14,186	25.4
モノレール	-	-	-	-
タクシー・ハイカー	10,117	10,087	1,993	3.3
レンタカー	2,243	3,432	1,364	0.7
その他の乗用車	9,337	4,790	3,198	3.0
旅客船・フェリー	-	-	-	-
国内線飛行機	5,788	2,091	322	1.9
その他の航空	9,421	4,127	1,307	3.0
不明	1,066	7,940	3,362	0.3
合計	306,609	277,171	46,168	100.0

注 「-」は手段として利用できない交通手段を示す。



仕様書等で作業内容は明示されていたか①

1. 当該作業内容の記載が仕様書等がない、又は不十分であったためプログラムミスが発生したと考えられるもの(5事案)

統計調査名	仕様書等記載との関係
①医療扶助実態調査(厚生労働省)	○受託事業者が電子レセプトデータを集計システムに取り込む際、電子レセプトの仕様及びその補足情報について認識を誤っていたため、プログラムの設計ミスにより、一部のデータを正しく取得できていなかった。
⑤健康保険・船員保険被保険者実態調査(厚生労働省)	○仕様書等において、数値の範囲に関する指示(標準報酬月額範囲)において計算式が用いられておらず、不明瞭だった。 ○その結果、受託事業者が、標準報酬月額等級等を確定するために設定する報酬月額範囲を指定する際の関数を間違えて設定(正「<」、誤「≤」)したことにより、統計数値に誤りを生じた。
⑥障害福祉サービス等従事者処遇状況等調査(厚生労働省)	○受託事業者がExcelで統計を作成する際に、本来集計範囲としない箇所までプログラム設定したことにより(正:福祉・介護職員処遇改善特別加算取得事業所の職員数、誤:福祉・介護職員処遇改善特別加算及び福祉・介護職員処遇改善加算(I~V)取得事業所の職員数)、統計数値に誤りを生じた。 ○総合評価方式による調達であり、調査手法、集計方法等は事業者の技術提案を求める項目であるため、仕様書等で詳細な作業方法は記載していない。作業内容の確認は、受託事業者との打ち合わせ、メール、電話等での協議により行なっていたが、プログラムミスが生じた作業について具体的な書面での指示は確認できなかった。

(→ 次頁に続く)

仕様書等で作業内容は明示されていたか①

(前頁の続き)

統計調査名	仕様書等記載との関係
⑦中小企業実態基本調査(中小企業庁)	○仕様書等において、外れ値の算出に使用する個票情報に関する注意点や、具体的な作業方法を指示しなかった結果、作業の過程において受託事業者が、対象外の個票を含めるとともに売上高に係る単位を誤って処理し、誤った外れ値を算出した。
⑨学術情報基盤実態調査(文部科学省) (記載すべき数値の設定に誤りがあった事案)	○受託事業者の作成した集計表テンプレートExcelファイルの書式設定が不適切であったため、0とすべき場所が空欄になる、記載されるべき数値が記載されないといった書式上の問題が生じた。 ○仕様書等には、「0」と空欄の区別が明確に記載されていなかった。

⇒ これらのケースにおいては、再発防止策として、仕様書等による作業内容の明確化が重要。

仕様書等で作業内容は明示されていたか②

2. 作業内容は仕様書等上、明確であったと考えられるもの(9事案)

統計調査名	仕様書等記載との関係
②情報通信業基本調査(総務省・経済産業省の共管)	<p>○外注事業者が業務効率化を図るために、集計表に係るシステムの見直しを行い、従来とは異なる集計プログラムを利用したところ、「卸売業」の合計値を算出するための設定が誤っていることに気付かないまま集計した。経済産業省担当者も当該箇所について従来とは異なる集計プログラムを使用していることを認識していなかった。</p> <p>○「卸売業」の合計値算出に係る仕様は明確に記載されていた。(従前と同様の定義、集計作業の指示)</p> <p>※審査・集計ツールの作成は事前に発注者の確認を得ることとなっていたが、受託事業者が自らの業務効率化のために行うシステム改修は仕様書等上報告を義務付けていない。</p>
③国勢調査(総務省)	<p>○「5年前常住市区町村「不詳」」の数値について、集計対象漏れにより過小値となっていたため、7表中1表を修正。市区町村×市区町村の長大な新規表であったため、統計センターでは、既存の審査システムで審査できず、別方法で審査を行ったがエラーが検出できなかった。</p> <p>○製表業務については、仕様書等に概要を提示している。「不詳」の項目は平成22年調査から追加されたものであり、定義は仕様書等に記載されていた。</p>
④通信利用動向調査(総務省)	<p>○集計項目は仕様書等上、明確であったが、集計用データを作成するプログラムを設計する際、データの引用箇所に指定誤りが生じたため、1～15の選択肢のうち、11～15の選択肢について、誤って再度1～5の回答データにて集計を行った。</p>

(→ 次頁に続く)

仕様書等で作業内容は明示されていたか②

(前頁の続き)

統計調査名	仕様書等記載との関係
⑧大学等におけるフルタイム換算データに関する調査(文部科学省)	<p>○集計項目(内数と合計の計算方法)は仕様書等上明確と考えられる。</p> <p>○A+Bの合計について、Bの内数を除去すべきところ、誤って二重に計上したものの。</p>
⑩訪日外国人消費動向調査(観光庁)	<p>○受託事業者が集計作業に使用しているプログラムにおいて、処理する変数を書き間違えたことにより、本来処理すべきでない変数によって処理が行われたため、該当する表5-1に一部修正が生じた。</p> <p>○仕様書等上、集計作業は明確に記載されており、前年度との変更点等詳細は口頭でも確認していた。</p>
<p>⑪学術情報基盤実態調査(文部科学省)</p> <p>⑫建設工事統計調査(建設工事施工統計調査)(国土交通省)</p> <p>⑬国際航空旅客動態調査(国土交通省)</p> <p>⑭全国貨物純流動調査(国土交通省)</p>	<p>○公表資料作成時のデータ参照先設定の誤り事案であり、仕様書等上、作業内容は十分明確と考えられる。</p>

⇒ これらのケースにおいては、作業内容は十分明確であったと考えられるため、

- ・ 受託事業者におけるチェック体制・方法
- ・ 発注者による履行確認

が再発防止に向け重要。

受託事業者への履行確認は行っていたか①(仮納品の報告)

- 作業内容は仕様書等上、明確であったと考えられる9事案について、履行確認の実施状況を分類。
- これら9事案のうち、5事案においては、一次納品(仮納品)としての報告をさせ、作業経過を確認していた。

仮納品の確認で誤りを防止できなかった理由は、

- ・前年度と比較して明確に誤りと認識できる数値ではなかった。
- ・納品物(PDF)の報告としては確認できない部分(プログラムによって参照される数値)の誤りまで発見できなかった。

などであった。

⇒ 一次納品(仮納品)としての報告をさせ、作業経過を確認するのみでは、目視確認で判明するような明確なミスは防いでも、一部の数値のみが誤っているといった事案を防ぐには限界があると考えられる。

受託事業者への履行確認は行っていたか②(テストデータによる検証)

仕様書等により、受託事業者にテストデータを用いた検証を義務付けることは、再発防止上、重要と考えられる。

○プログラム作成時や変更時のテストデータを用いた検証については、9事案中7事案で行われていなかった。

○テストデータを用いた検証を行っていたが誤りを防止できなかった2事案については、
・長大な新規表が対象の事案であり、別方法で審査を行っていた。
・改修箇所について重点的に確認を行っていた。
ことによるものであり、例外的なケースと考えられる。

統計調査名	履行確認の状況
③国勢調査(総務省)	○疑似データを使用し、仕様どおり作られているかを確認する単体テストと、全てのプログラムをつなげて行う総合テストを実施している。 ○今回の事案では、市区町村×市区町村の長大な新規表であったため、統計センターでは、既存の審査システムで審査できず、別方法で審査を行ったがエラーが検出できなかった。
⑪学術情報基盤実態調査(文部科学省)(プログラム作成時のマッピングのずれが生じた事案)	○調査担当者は、テスト用のデータもしくは当該年度に回収した回答データを用いて、プログラムから出力された集計結果(=公表用の統計表)と、調査担当者がプログラムを用いずに独自に集計した結果との突き合わせを行い、特に改修のあった部分を中心にプログラムが正常に動作するかどうかを確認している。 ○受託事業者は、毎年度変更点についてプログラムの改修をしており、受託事業者及び調査担当者が動作テストを行う際には改修した箇所を重点的に確認するため、変更がなかった箇所における不具合が発見されなかったと考えられる。

(参考)追加調査での確認事項

第6回点検検証部会(令和元年6月13日)での審議を経て、基幹統計、一般統計を所管している府省に対し、書面調査や一斉点検で報告のあった結果数値の訂正案件について、プログラムミス(プログラムの運用面でのミスを含む)が原因で発生したものを追加調査にて確認を行なった。

項目	確認事項(各府省の記入内容)
統計名	当該統計名
府省庁等名 (担当課室名)	当該統計を所管する府省庁等名(担当課室名)
統計の種類 (基幹/一般)	基幹統計、一般統計の別
事例報告の区分	1:書面調査で報告された結果数値訂正事案 2:一般統計自己点検結果における結果数値訂正事案
公表時期	結果数値訂正事案の公表時期
プログラムミス (プログラムの 運用面のミス を含む)での誤 りの有無	プログラムに起因する数値訂正か否か (※プログラムに起因する数値訂正であった場合は、以下 を確認)
発生した作業 工程	1:個票処理過程 2:集計処理過程 3:公表処理過程(集計処理終了から公表までの間) 4:その他(具体説明)
発見した契機	具体説明(例:前月比を確認した際、振れ幅が大きかった ため など)

項目	確認事項(各府省の記入内容)
事象の区分	1:プログラムの設計上に誤りがあった。(発注者側の指示 誤りを含む。) 2:プログラム内部に誤り(バグ)があった。 3:その他(プログラムそのものに不備はないが、運用面 での誤りとして、データの投入場所の誤りや作業手順の誤り 等の人為的ミスがあった場合)
事象の具体	内容、原因を具体的に説明(例:××、○○に係る個票 データをプログラムに取り込む際に付番する県コードの誤 りで、○表中○表を修正 などバグ等の内容や結果への 影響の状況、○○集計表作成における数式入力の際の参 照箇所のミス、等について具体的に記述。また、業者との 間でのコミュニケーションエラーが関連している場合には、 具体的内容とともにその旨を付記。)
履行確認の実 施状況有無	1:(プログラム作成時や変更時)テストデータ等を用いた検 証を行ったか 2:作業で発生したデータのエラー件数を報告させているか 3:一次納品(仮納品)としての報告をさせて、作業経過を確 認しているか
テストデータ等 の具体	(上記で1と回答した場合)テストデータ等の内容を具体的 に説明(どのようなテストデータを誰が(府省側か業者側 か)作成、確認したのか など)