データ作成マニュアル作成の

ポイント

令和６年

国土交通省　都市計画課　都市計画調査室

**はじめに　データ作成マニュアル作成のポイントの目的**

郵送回収された調査票のデータ入力およびWEB回答のダウンロードにより回答データが作成されます。回答データは、コンピュータシステムによる最終的なチェックと修正を経て、オリジナルファイルとなります。最後に、オリジナルファイルを集計して、有効世帯数、有効個人数、トリップ数などの回収状況を整理します。なお、オリジナルファイルに拡大処理を施したものをマスターファイルと呼びますが、拡大処理は本マニュアル作成のポイントの対象外としています。本マニュアル作成のポイントでは、これらの作業①～④の標準的な手順、内容を示しています。

なお、本資料は、国が各都市圏での円滑な調査実施を支援するために、パーソントリップ調査の回答データの作成についてのポイントとマニュアルの一例を示したものです。調査の実施にあたっては、地域の実情に合わせて適宜、適切な内容に修正し、活用ください。

エディティング・コーディングが終了した

郵送回収の「世帯票」、「個人票」

①　回答データの作成

・郵送回収調査票のデータ入力

・WEB回答データのダウンロードと処理

②　システムチェック

・入力ロット別のチェック

・全体データでのチェック

③　オリジナルファイルの作成

＊システムチェック完了後のデータが

オリジナルファイルになります。

④　調査結果の集計

・オリジナルファイルの集計

・無効内訳の整理

・ データ修正

照合データ

○回収状況に関するデータ

○内容に関するデータ

　　調査日一覧、コード表等

マスターファイルの作成

**〈調査の準備〉**

**○対象世帯名簿データの作成**

**○調査配布物の作成**

「抽出・対象世帯名簿整理マニュアル作成のポイント」を参照

**〈調査の実施〉**

**○調査票等の発送**

・発送スケジュール

・発送物件の作成

・発送・返送の手段

・発送戻りの処理

**○問合せへの応対**

・対象者等からの問合せ

（電話、WEBフォーム）

**○郵送回収調査票の整理**

・返送封筒の受取・仕分け整理

・返送封筒の開封とバッチ作成

**〈データ整備〉**

**○エディティング**

・調査票の点検、修正

**○コーディング**

・住所等のコード番号記入

**○郵送回収調査票のデータ入力**

**○WEB回答データの処理**

**○データチェック・修正（WEB回答含む）**

**○オリジナルファイルの作成（郵送＋WEB）**

**○有効回収率等の集計**

・有効回収率／無効内訳

・トリップ状況（原単位等）

**〈進捗管理〉**

**○進捗状況の管理**

・進捗状況の整理と報告

「調査実施・管理マニュアル作成のポイント」を参照

「調査実施・管理マニュアル作成のポイント」を参照

「エディティングマニュアル作成のポイント」を参照

「コーディングマニュアル作成のポイント」を参照

**本資料**

**「データ作成マニュアル作成のポイント」の記載範囲**

図　本マニュアル作成のポイントの対象範囲

データ作成マニュアル

令和○年

○○市　○○課

目次

[1章. 回答データの作成 1](#_Toc167464779)

[1. 回答データの作成方法 2](#_Toc167464780)

[2. 個人情報の取扱いについて 2](#_Toc167464781)

[3. 調査票のデータ入力 3](#_Toc167464782)

[4. WEB回答のデータ出力 3](#_Toc167464783)

[5. 回答データのデータ構成 4](#_Toc167464784)

[6. 回答データのデータレイアウト 5](#_Toc167464785)

[2章. システムチェック 9](#_Toc167464786)

[1. システムチェックの作業手順 10](#_Toc167464787)

[2. チェックのレベル 10](#_Toc167464788)

[3. 外部データとの照合 11](#_Toc167464789)

[4. チェック項目 13](#_Toc167464790)

[5. エラー修正 13](#_Toc167464791)

[3章. オリジナルファイルの作成 18](#_Toc167464792)

[4章. 調査結果の集計 20](#_Toc167464793)

[1. オリジナルファイルの集計 21](#_Toc167464794)

[2. 無効内訳の整理 21](#_Toc167464795)

#

# 回答データの作成

**考え方**

データ作成にあたっては、まずは、コーディング後のデータ入力により回答データを作成します。本章は、回答データの作成について、データの形式に応じた作成手順や回答データのレイアウト等を説明・整理しています。

## 回答データの作成方法

郵送回収の調査票は、エディティング、コーディングの後、6ページから8ページに示す所定のレイアウトに沿ってデータ入力を行います。

WEB調査によりサーバー（推進本部が設置）に蓄積されたWEB回答データはダウンロードした上で、上記の所定のレイアウトに合うように、各サポートセンター（実施本部）においてデータレイアウト変換等の処理を加えます。

## 個人情報の取扱いについて

平成17年4月より個人情報保護法が施行され、個人情報に関する世論の関心も非常に高まっています。調査対象者名簿や調査配布物、電話応対メモについては、個人情報保護法等関係法令にもとづいて、適切に扱うことが必要となります。

データ作成作業に当っては、個人情報の管理に十分心がけ、紛失、漏洩、毀損等の事故を発生させないように、十分な配慮を持って対応してください。

（該当する場合のみ）なお、三大都市圏での調査（一般統計調査）、都道府県および政令指定都市が主体となる調査（届出調査）の調査票情報は、統計法に基づき、適切に扱うことが必要になります。

注）調査主体に応じて、適切に設定してください。

＊個人情報とは

* 個人情報とは、「氏名、生年月日その他の記述等により特定の個人を識別することができる個人に関する情報」、又は「個人識別符号が含まれるもの」（個人情報保護委員会発行の個人情報の保護に関する法律についてのガイドラインより）のことを指します。
* 個人認識符号とは以下の通り（個人情報の保護に関する法律　第2条第2項）。

①特定の個人の身体の一部の特徴を電子計算機の用に供するために変換した文字、番号、記号その他の符号であって、当該特定の個人を識別することができるもの

②個人に提供される役務の利用若しくは個人に販売される商品の購入に関し割り当てられ、又は個人に発行されるカードその他の書類に記載され、若しくは電磁的方式により記録された文字、番号、記号その他の符号であって、その利用者若しくは購入者又は発行を受ける者ごとに異なるものとなるように割り当てられ、又は記載され、若しくは記録されることにより、特定の利用者若しくは購入者又は発行を受ける者を識別することができるもの

#### 取扱に注意する物件

対象者が特定される物件は、個人情報の観点から取扱に注意してください。

#### 具体的な注意点

①　取扱に注意する物件を取り扱う作業は、原則としてサポートセンター（実施本部）内で行います。

②　取扱に注意する物件は、複写・複製を厳禁とします。

③　万一、取扱に注意する物件について、紛失、漏洩、毀損等の事故が発生した場合には、直ちに管理スタッフ（受託会社の社員）に連絡します。

④　作業中に知り得た対象世帯構成状況等のプライバシー情報は、作業中・作業終了後に口外してはいけません。

## 調査票のデータ入力

①　データ入力の対象

* 有効回収世帯を対象に、バッチの並び、バッチ内の世帯の並びのままに、下記のデータ入力を行います。

世帯票／個人票

②　データ入力の手順

* 入力作業は、作業の効率性や安全性等を考慮して、適切な回数に分割して行います。
* 入力会社に依頼する場合には、受渡し・返却の都度、個人情報を授受したことがわかる確認証を取り交わします。

また、下記の「調査票入力記録表」に記録します。

■調査票入力記録表（標準例）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 入力ロットNo. | 調査対象都市名 | バッチ番号 | 入力会社への受渡日 | 入力会社　からの納品日 | 入力会社名（所在地） |
| 調査主体 |
|  |  |  | ○○～○○ |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

## WEB回答のデータ出力

WEB回答のデータについて出力し、6ページから8ページに示す所定のレイアウトにあわせる加工を行います。

## 回答データのデータ構成

有効回収世帯について、世帯票、および個人票の各データから構成します。

■データ構成イメージ

2番目個人の個人票・第1レコード

世帯1

世帯票・1番目の個人

世帯票・2番目の個人

1番目個人の個人票・第1レコード

　＊第1トリップ到着地等まで含む

1番目個人の個人票・第2レコード

世帯2

入力ロット①

△△△市

バッチ1

バッチ2

□□□市

入力ロット②

入力ロット③

WEB回答

■票区分

|  |  |
| --- | --- |
| 種類 | 備考 |
| 有効世帯 | 世帯票 | 世帯人員ごとに1レコード＊レコード数は「世帯人数（5歳未満含む）」と一致する |
| 有効な個人票 | トリップごとに1レコード＊「トリップなし」の場合、第1レコードのみで、「はじめにいた場所」までのデータ＊「トリップあり」の場合、個人ごとのレコード数はトリップ数と一致する |

## 回答データのデータレイアウト

回答データのレイアウトについては、本章の次頁以降に示すデータレイアウトを活用することを推奨します。

■世帯票



■個人票　1/2



■個人票　2/2



# システムチェック

**考え方**

本章は、作成された回答データに対して、コンピュータシステムを用いてエラーチェックを行うにあたって、システムの構造や流れ、一つ一つのチェック項目の考え方等を整理しています。

## システムチェックの作業手順

システムチェックは、サポートセンター（実施本部）スタッフによるチェック（エディティング）を経たデータを、コンピュータによって最終チェックする工程です。システムチェックによりエラーが検出された箇所は調査票（原票）と照らし合わせて修正し、エラーが検出されなくなるまで、入力データの修正とチェックの作業を繰り返します。

チェック項目は、以下の内容から構成されています。

* プレチェック　；票区分等によるデータ構成
* キー部チェック；整理番号等による調査対象者の重複、漏れ
* 内容チェック　；回答内容の記入もれや矛盾

入力データ

＊破線内は各ロット内でそれぞれ処理

システムチェック

・プレチェック

・キー部チェック

・内容チェック

修正データの入力

エラー修正

・調査票（原票）との照合

エラーリストの出力

エラー

有無

終了

オリジナルファイルの作成

図　システムチェックの作業手順

## チェックのレベル

チェックレベルは、「異常レベル」と「警告レベル」の2種類に分けられています。

表　チェックのレベル

|  |  |
| --- | --- |
| チェックレベル | 処理方法 |
| 異常レベル | 必ず修正を要するエラー |
| 警告レベル | 必ずしも修正を必要としないが原票と照合して確認を要するエラー |

## 外部データとの照合

#### 照合データの種類

エラーチェックシステムにおけるチェック項目は、以下の3種類に大きく分けられます。

* 単一の入力データのチェック

（例：ある設問の回答が、指定された範囲内にあるか）

* 複数の入力データ間のチェック

（例：ある設問の回答が、他の設問の回答と矛盾しないか）

* 入力データと外部データを照合して行うチェック

（例：ある設問の回答が、外部データのリストに含まれているか）

世帯票

個人票

**一致（整理番号）**

**一致（整理番号）**

**整合**

回収管理データ

調査日一覧

内容に関するデータ

・ゾーンコード表

・鉄道駅コード表

・駅勢圏テーブル（あれば）

**整合**

図　外部データとの照合イメージ

「入力データと外部データを照合して行うチェック」において入力データと照合する外部データは、下表の通りです。

表　照合する外部データ一覧

|  |  |
| --- | --- |
| 照合データ | 内容 |
| 回収管理データ | 整理番号等 |
| 調査日一覧 | 指定した調査日（都県市ごと） |
| ゾーンコード表 | 地区ゾーンを示すコード |
| 鉄道駅コード表 | 鉄道駅を示すコード |
| 駅勢圏テーブル（あれば） | 駅・利用想定ゾーンの対応 |

#### 回収管理データ（システムチェック用）の作成と更新

回収管理データは、回収時に作成され、エディティングで更新されたデータにもとづいて作成されます（「調査実施・管理マニュアル」および「エディティングマニュアル」を参照）。

回収管理データ（システムチェック用）は、このうち有効世帯（回収管理データで「受取日」の記載があり、かつ「無効世帯の無効理由」がブランクである世帯）を対象に、整理番号の列を抽出し、作成します。

システムチェックでは、この回収管理データ（システムチェック用）と回答データの整理番号の照合を行います。システムチェック作業の中で、回収管理データ（システムチェック用）に修正が生じた場合には、更新した上でシステムチェックを繰り返すことになります。

## チェック項目

エラーチェックは、次ページ以降の表に示す視点にもとづいて行います。

## エラー修正

システムチェックで検出されたエラー（異常・警告）については、調査票（原票）と照合の上、データ修正を行います。

■世帯票のチェック項目

| 項目 | 視点 |
| --- | --- |
| 回収分類 | * 選択肢の範囲にあるか
 |
| 世帯番号 | * 正しく一意に世帯を特定できるか
 |
| 世帯人数\_5歳未満含む | * 世帯票の「個人レコード数」と一致するか
 |
| 世帯人数\_5歳未満除く | * 世帯票の5歳未満を除く「個人レコード数」と一致するか
 |
| 現住所\_ゾーンコード | * ゾーンコード表にあるか
 |
| 現住所\_町字ID | * 町字コード表にあるか
 |
| 現住所\_地番ID現住所\_街区ID現住所\_住居ID現住所\_住居2ID | * 矛盾なく住所が記入されているか
 |
| 自動車台数 | * 数字が大き過ぎず常識の範囲内であるか
* 「自由に使える自動車」において「ほぼ自分専用の自動車がある」と回答した人数と矛盾していないか
 |
| 世帯年収 | * 選択肢の範囲にあるか
 |
| 世帯内番号 | * 正しく一意に個人を特定できるか
 |
| 性別 | * 選択肢の範囲にあるか
 |
| 年齢 | * 数字が大き過ぎず常識の範囲内であるか
* 続柄との関係に矛盾がないか
 |
| 世帯主との続柄 | * 選択肢の範囲にあるか
* 世帯内の第１レコードが世帯主本人であるか
 |
| 就業形態 | * 選択肢の範囲にあるか
* 年齢との関係に矛盾がないか
 |
| 職業 | * 選択肢の範囲にあるか
* 就業形態との関係に矛盾がないか
* 年齢との関係に矛盾がないか
 |
| 産業 | * 選択肢の範囲にあるか
* 就業形態との関係に矛盾がないか
 |
| 自動車運転免許 | * 選択肢の範囲にあるか
* 年齢との関係に矛盾がないか
 |
| 自由に使える自動車 | * 選択肢の範囲にあるか
* 自動車運転免許との関係に矛盾がないか
* 年齢との関係に矛盾がないか
* 自動車台数との関係に矛盾がないか
 |
| 外出に関する身体的な困難さ | * 選択肢の範囲にあるか
 |
| 勤務先・通学先・通園先\_ゾーンコード | * ゾーンコード表にあるか
* 就業形態との関係に矛盾がないか
 |
| 勤務先・通学先・通園先\_町字ID | * 町字コード表にあるか
* 就業形態との関係に矛盾がないか
 |
| 勤務先・通学先・通園先\_地番ID勤務先・通学先・通園先\_街区ID勤務先・通学先・通園先\_住居ID勤務先・通学先・通園先\_住居2ID | * 矛盾なく住所が記入されているか
* 就業形態との関係に矛盾がないか
 |

■個人票のチェック項目

| 項目 | 視点 |
| --- | --- |
| 回収分類 | * 選択肢の範囲にあるか
 |
| 世帯番号 | * 世帯票と紐づいているか
 |
| 世帯内番号 | * 世帯票と紐づいているか
 |
| 調査日 | * 指定調査日の範囲内にあるか
 |
| 平日休日 | * 選択肢の範囲にあるか
 |
| トリップ有無 | * 選択肢の範囲にあるか
* トリップ情報の回答状況との関係に矛盾がないか
 |
| トリップ数 | * 数字が大き過ぎず常識の範囲内であるか
* トリップ情報の回答状況との関係に矛盾がないか
 |
| トリップ番号 | * 第1トリップから連番になっているか
 |
| 出発地\_区分 | * 選択肢の範囲にあるか
* 第1トリップのときのみ記入されているか
* 区分が勤務先・通学先・通園先と回答されているとき、世帯票の勤務先・通学先・通園先に回答されているか
 |
| 出発地\_ゾーンコード | * ゾーンコード表にあるか
* 出発地\_区分との関係に矛盾がないか
 |
| 出発地\_町字ID | * 町字コード表にあるか
* 出発地\_区分との関係に矛盾がないか
 |
| 出発地\_地番ID出発地\_街区ID出発地\_住居ID出発地\_住居2ID | * 矛盾なく住所が記入されているか
* 出発地\_区分との関係に矛盾がないか
 |
| 到着地\_区分 | * 選択肢の範囲にあるか
* 自宅へのトリップや勤務先・通学先・通園先へのトリップが連続していないか
* 区分が勤務先・通学先・通園先と回答されているとき、世帯票の勤務先・通学先・通園先に回答されているか
* 目的との関係に矛盾がないか
 |
| 到着地\_ゾーンコード | * ゾーンコード表にあるか
* 到着地\_区分との関係に矛盾がないか
 |
| 到着地\_町字ID | * 町字コード表にあるか
* 到着地\_区分との関係に矛盾がないか
 |
| 到着地\_地番ID到着地\_街区ID到着地\_住居ID到着地\_住居2ID | * 矛盾なく住所が記入されているか
* 出発地\_区分との関係に矛盾がないか
 |
| 出発時刻 | * 時刻として矛盾がない値になっているか
* 直前トリップの到着時刻との関係に矛盾がないか
 |
| 到着時刻 | * 時刻として矛盾がない値になっているか
* 出発時刻との関係に矛盾がないか
* 1トリップのトリップ時間が長過ぎず常識の範囲内であるか
* 1日のトリップ時間の合計が24時間以内か
 |
| 目的 | * 選択肢の範囲にあるか
* 年齢との関係に矛盾がないか
* 到着地\_区分との関係に矛盾がないか
* 就業形態との関係に矛盾がないか
* 勤務先での仕事目的のトリップ、学業目的のトリップや帰宅目的のトリップが連続していないか
 |
| 交通手段 | * 選択肢の範囲にあるか
* 同じ交通手段が連続していないか
* 1番目の交通手段から順番に記入されているか
* 鉄道、路線バス・コミュニティバス、船舶・フェリー、航空機、その他の交通手段でトリップが始まっていないか
* 鉄道、路線バス・コミュニティバス、船舶・フェリー、航空機、その他の交通手段でトリップが終わっていないか
 |
| 乗換地点 | * 鉄道駅コード表にあるか
* 鉄道を利用した時に回答されているか
* 出発地や到着地が乗車駅や降車駅の駅利用圏にあるか
* 乗車駅と降車駅が同じでないか
 |
| 運転有無 | * 選択肢の範囲にあるか
* 自動車を利用した時に回答されているか
* 自動車運転免許との関係に矛盾がないか
 |

# オリジナルファイルの作成

**考え方**

システムチェック後に、オリジナルファイルを作成します。本章は、オリジナルファイルを作成するにあたっての処理方法を整理しています。

入力ロット別データのシステムチェックが全て終了した後、データ全体を統合して、オリジナルファイルを作成します。具体的には、以下の処理を行います。

①　入力ロット別のシステムチェックの終了

「異常」エラーの件数がなくなった状態で、個々のロットについてのデータシステムチェックを終了します。

②　全体データのシステムチェック

全体データを対象に前記のシステムチェックを行います。全体データのチェックは、世帯や個人でのデータ重複の検出が主なねらいですので、回答内容に関する「警告」エラーのチェック項目をはずすなど、原則として「異常」エラーに関するチェックにしても構いません。

③　オリジナルファイルの作成

全体データのシステムチェックを完了し、そのデータをもってオリジナルファイルとします。なお、オリジナルファイルのレイアウトは、入力データと同じとします。

# 調査結果の集計

**考え方**

作成されたオリジナルファイルを集計し、回答状況を確認することが望ましいといえます。本章は、オリジナルファイルを集計し、回収状況を把握するための考え方を整理・説明しています。集計の視点は、回収状況、記入状況、トリップ特性等が考えられます。

## オリジナルファイルの集計

各市区町村における有効回収状況を把握するため、市区町村別の有効回収率・達成率を集計して、有効回収の目標（目安）とした総数が確保されていること、市区町村別にも一定の水準（発注者と協議）にあることを確認します。

＜集計の視点＞

○回収状況

－有効回収数

－有効回収率

○記入状況

－各調査項目の不明率

　 ※現住所、勤務先・通学先・通園先、はじめにいた場所、到着地に関しては、番地・号が特定されているデータの割合も確認

○トリップ特性

－トリップ数

－生成原単位（ネット及びグロス）

－発生量、集中量及び発生集中量

－各交通手段の分担率

－発時刻、着時刻及び所要時間

 ※必要に応じて、目的別、性別・年齢別に上記を集計しチェック

## 無効内訳の整理

調査票発送からオリジナルファイル作成の全過程で生じた無効の状況を最終的に整理します。

①　回収されたが無効の世帯（回収無効世帯）

回収管理データにもとづき整理します。

②　世帯単位での無効の全体内訳（全体無効内訳）

調査対象世帯数と有効回収世帯数の差で無効世帯の合計値を先に算定します。無効内訳のうち、調査票発送戻り世帯数、電話拒否世帯数および回収無効世帯数は従前に整理した値として、残りの差分を無回答として整理します。

|  |  |  |
| --- | --- | --- |
| 改訂履歴一覧 |  |  |
|  | マニュアル名： | データ作成マニュアル |
|  |  |  |  |
| 整理番号 | 日付 | 改訂箇所 | 改訂内容 |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |
| 6 |  |  |  |
| 7 |  |  |  |
| 8 |  |  |  |
| 9 |  |  |  |
| 10 |  |  |  |
| 11 |  |  |  |
| 12 |  |  |  |
| 13 |  |  |  |
| 14 |  |  |  |