

## 第2回 データ回収から入力まで (4/17)

作成：田中重人 (講師) <tsigeto@nik.sal.tohoku.ac.jp>

### 1 回収直後の作業

#### 1.1 ID 番号の付与

調査票上部などに通し番号を振る。以降は、この ID 番号によって回答者を識別する。

#### 1.2 エディティング (editing)

記入漏れ、不完全な回答、誤りなどを探して訂正する。可能であれば、再調査をおこなうこともある。訂正する際には、回答とは区別できるように色ペンなどを使う。

論理的にありえない回答の組み合わせや、珍しい回答についても注意する。

記入漏れの非常に多いもの、内容の信頼性が低いものは無効とする。

有効回収率 = 有効票数 / 配布票数

### 2 コード表

コンピュータでの分析のためには、データはすべて数値で表すのが便利である。

回答と数値との対応表を「コード表」という。

コード表は、調査票作成と同時に作っておくとよい。ただし、調査後に自由回答のアフターコードなどを追加した場合には、コード表のほうも改訂する必要がある。

欠損値 (無回答・非該当など) にどのような数値を与えるか?

### 3 コーディング

コード表にしたがって回答を数値化していく作業。

- 調査票に直接書き込む (色ペンで回答と区別する)
- 専用の用紙に転記する

このコード結果を見ながらコンピュータに打ち込むことになる。

## 4 コンピュータファイルの作成

変数名を決めておく

- 英字と数字のみ (A-Z 0-9 \_)。ただし最初の1文字は数字を使えない
- 大文字 / 小文字を区別しない
- 8文字以内
- 問番号にしたがって、Q10\_4A などとつけると便利 (項目内容については、SPSSで「ラベル」をつけるとよい)
- 転記用紙を使うときは、最初から変数名を書き込んでおくとうよい

コンピュータに打ち込む。SPSS 等で作成してもよいが、Excel 等の表計算ソフトが使いやすい。Excel を使う場合の注意：

- 最初の行に変数名を書き込んでおく
- 「ウインドウ枠を固定」しておくとう便利
- 別の人に打ち込んでもらって、ファイルの相違をチェックするのがよい
- SPSS で分析するときは、ファイルの種類で「xls」を指定すると、そのまま読み込める。オプションで、1行目に変数名であることを指定すること

## 5 クリーニング

データにみられるまちがいを摘発・修正する

- 度数分布表やクロス表を出力して、ありえない値をチェックする
- 変数の値の範囲や、変数間の依存関係をあらかじめリストアップしておく (コード表に書き込んでおくとうよい)

あとで分析を進めていくうちにまちがいがみつかることもある。