

SPSSによるクロス集計 操作のキソのキソ

村瀬 洋一

2007/12

大まかな手順

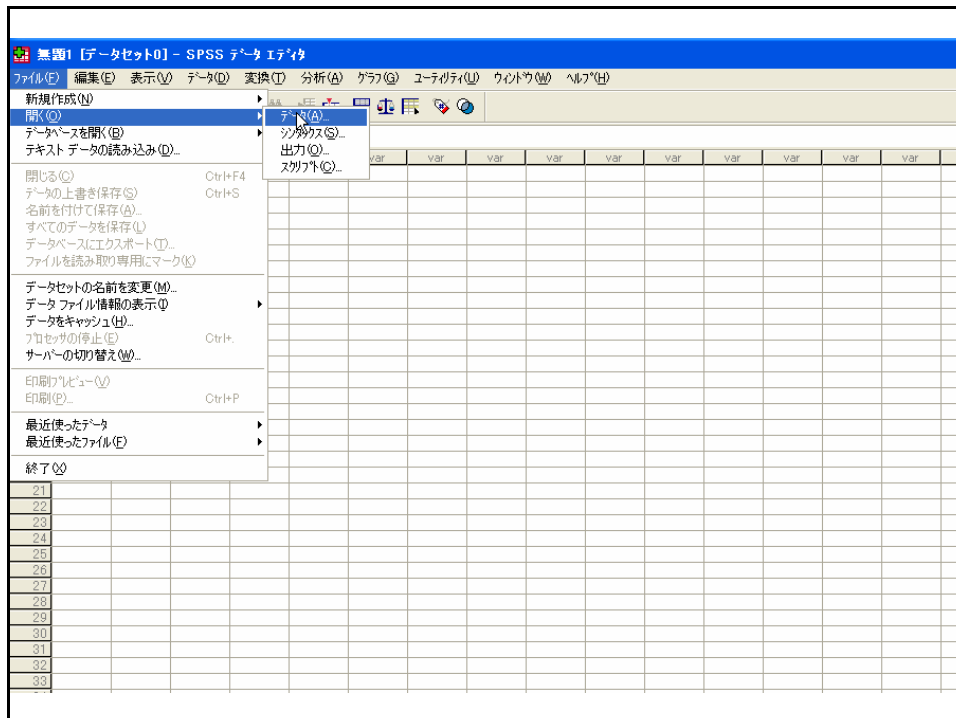
1. まずデータを自分のハードディスクに保存
2. その後分析用ソフトSPSSを起動して分析
3. 分析結果の表は、エクセルで線を引くなど編集

1. 分析用データの保存

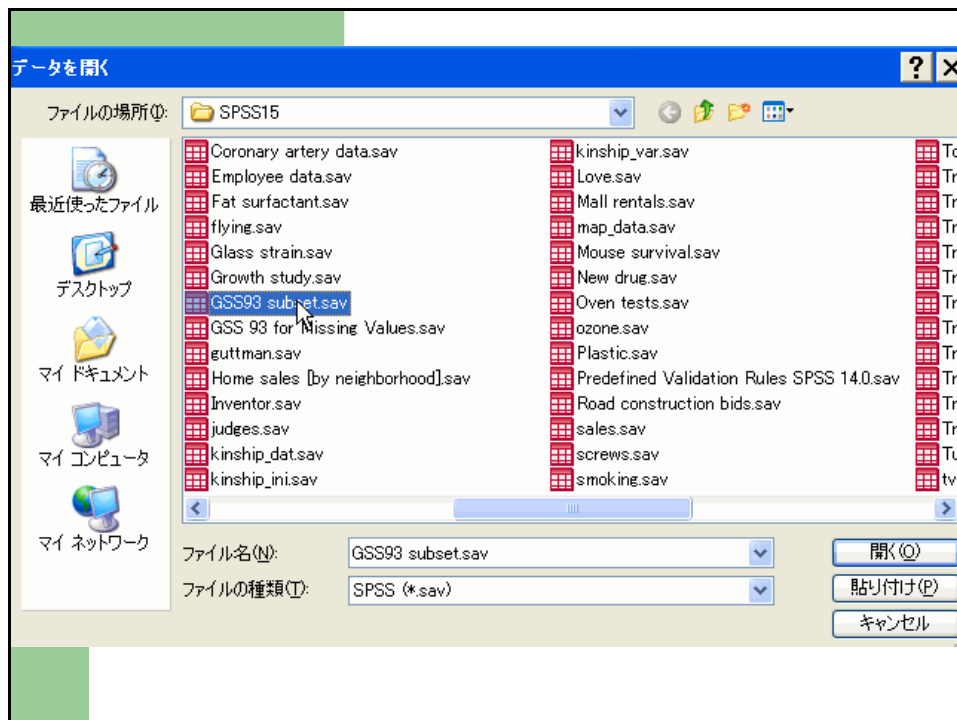
- まずデータを入手。データベースで検索するか、この科目ホームページの、[資料その3](#)などを見てください(自分の調査データやSPSSフォルダにあるファイルも可)。
- データは各種あります。まずは [SPSS形式データファイル](#) (拡張子がsavのもの) を選ぶと簡単。データのリンクを右クリックすると保存できます。
- 保存先は自分のマイドキュメント(立教ではホームディレクトリ、Hドライブと同じ)にしてください。立教のシステムでは、自分のファイルはHドライブに保存します。

2.SPSSの起動

- スタートボタンを押してSPSSを起動し、画面の上にある「ファイル」をクリックして、先ほど保存したデータファイルを開く
なお、最初に出てくる「チュートリアル実行」画面はキャンセルボタンを押して消せばよい。
- データファイルをダブルクリックしてSPSSを起動してもよいが、エラー表示が隠れていて動かないことがある。その場合、隠れているエラーのOKボタンを押してから、SPSS画面上の「ファイル」をクリックして「開く」を選びデータファイルを開くこと
- ホームページ上で「開く」をクリックしても動かないことがあるので、必ずスタートボタンを押してSPSSを起動してから、上記のように操作する。



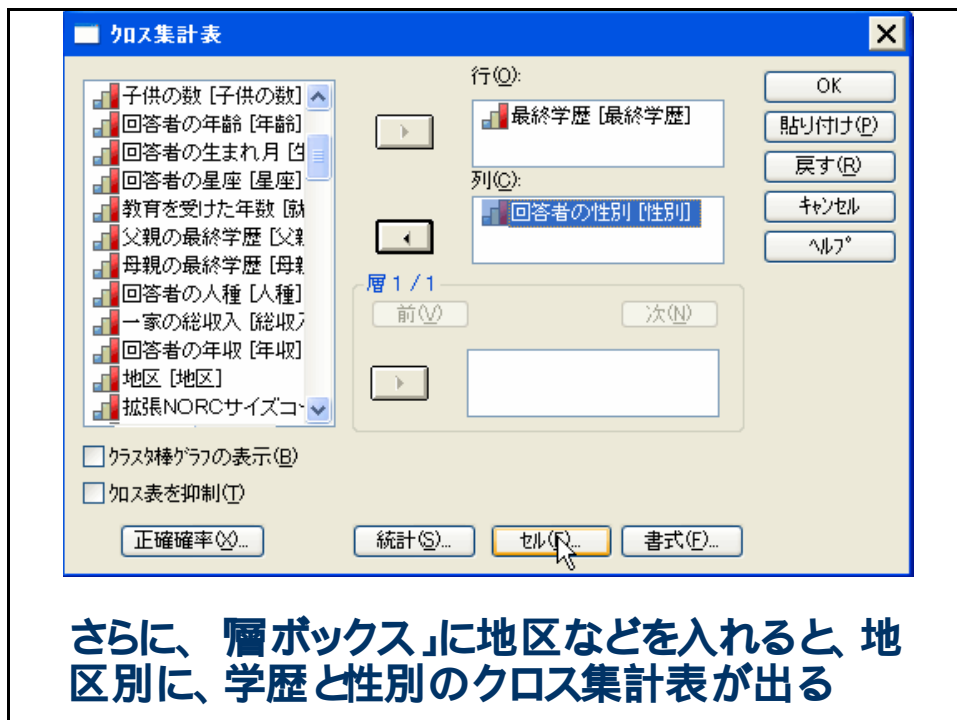
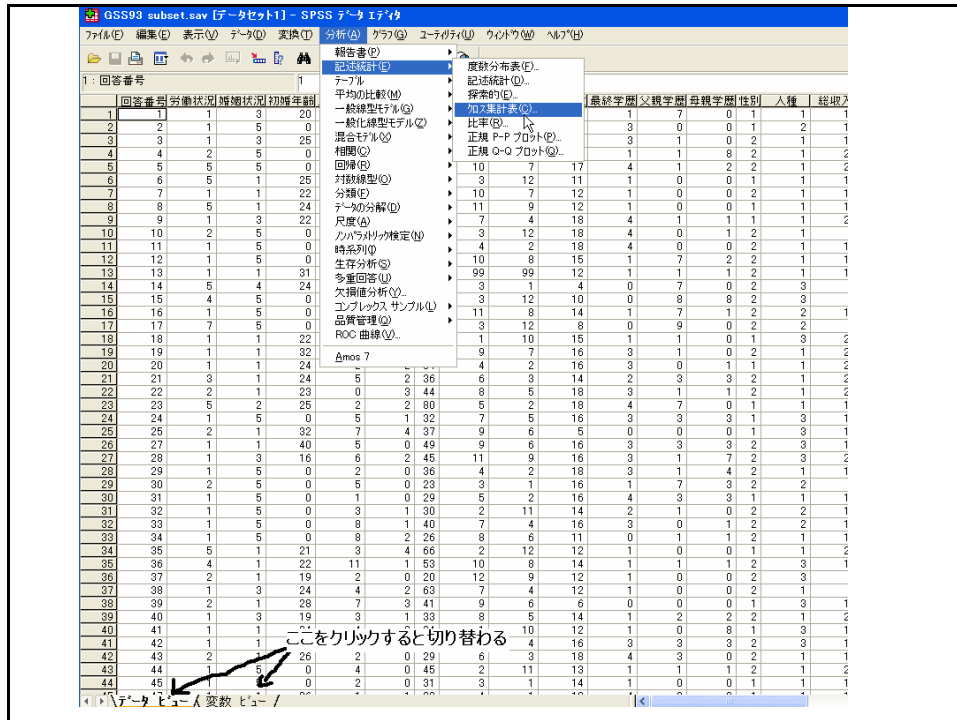
- まずは、「SPSSフォルダ」の中にあるGSS93subsetという練習用データを使ってもよい(次画面参照)。
- これはアメリカの社会調査の結果。SPSSフォルダに練習用に入ってるファイルです。おもしろい問が多い。好きな問を選んで自由に分析しましょう。
- データビューの画面では、横1列が1人の回答なので、1000人分なら1000行あることを理解しましょう。それが分かれば簡単。



3. シンタックスを書く

- SPSS画面上の「ファイル」をクリックして、「新規作成」 「シンタックス」を選ぶ。
- クロス集計用のシンタックスを書いてください(後述の7.にある見本3行のみでよい)。
- V29など変数名は、好きな変数名に書き換えてください。ただし、存在する変数名でなくてはならない。変数名は変数ビューで確認できる。

- 次ページのようにシンタックスは書かず、画面の上にある「分析」をクリックしクロス集計を選んで同じ結果になる。まずは自分が分析したい好きな変数を行、性別などを列ボックスに入れる。
- 次に、「セル」ボタンを押して「列%」(または行%)のみを出す設定にして、OKボタンを押す。
度数(人数)のチェックをはずした方が分かりやすい表になる。%と人数が混ざると見にくい。ただし、数人しかないセルもありうるので分布に注意。



4.変数名の確認

- [SPSSのデータ画面](#)を見て、変数名を確認してください。画面下にある「[変数ビュー](#)」をクリックすると、変数の詳しい説明が表示されます。
- 変数名はQ3やV3など、データによって違います。
- [調査票のファイル](#) (PDFやワードファイルなど)を見れば、各問の詳しい内容が分かります。自分でよく見るのが重要。好きな問を選んで分析。

5.シンタックス実行

- [シンタックスを実行](#) (ボタンか、画面上の「実行」をクリック)すればクロス集計表を作成できます。ピリオドを最後に1つだけつけるなど、忘れずに。
- [エラーが出る場合「シンタックスの基本解説PDF」](#)を見て、エラー内容を確認し、シンタックスを修正。
- 変数名が間違っている時や、シンタックスに全角スペースが入っている時などにエラーが出ます。

6.分析の手順

- まず、自分が分析したい問 (被説明変数 Y になるもの) を1つ決める。
- その上で、性別、年齢別、学歴別などの基本的なクロス表を出してみると良い (基本属性との関連を確認)。
- 次に、基本的な表の他、Yの原因となっていると考えられる問 (説明変数 X となるもの) について分析します。X が基本属性でもよいでしょう。
- なお、必要に応じてカテゴリー合併を行うこと。

7.シンタックス見本

- クロス集計シンタックスで、変数を指定するTAB行に V3 by V29 by 性別変数名 を書くと、男女別にクロス集計結果が出ます。

クロス集計のシンタックス見本

CRO

/TAB = V3 by V29

/CEL = ROW .

- ◆ ROW の代わりにCOLと書くと縦%が出ます。
- ◆ ROW の代わりにCOUNTと書くと実数が出ます。

8. カテゴリー合併

- クロス集計をする時、各変数が細かすぎると分かりにくい。例えば年齢が数十あると表としては良くない。低、中、高年齢の3カテゴリー等に直す。
- 各変数のカテゴリーは多くても5くらいにすること。
- 元の変数が細かすぎる場合、必ずクロス集計の前にカテゴリー合併を行う。
- カテゴリー合併は、別資料「SPSSシンタックスによる分析」をよく読み、シンタックスを用いるとよい。

9. 結果のまとめ方

- 結果の表をマウスで選択し、コピーしてエクセルなどに貼り付けることができます (あるいは「ファイル」をクリックしエクセル形式でエクスポート可)。
- **クロス集計表**は、別の資料「表とグラフ形式について」を参考に、横線のみ、表中には横%のみ (あるいは縦%のみ)を書き、合計の所にだけ人数を書いた形式に。
- エクセルで、セルを右クリックして書式設定を選ぶと、表示形式や罫線の変更ができます。%は小数点以下は四捨五入した表示で。グラフにするとおよい。

10. データの内容を絞る場合

- 例えば、多くの国を含んだデータで、日本のデータのみをしたい場合、
変数V3で日本が24 なら、SPSSのシンタックス画面で、

```
select if v3 =24.
```


と書いて実行。すると、v3が24の人だけにデータ人数が小さくなる（それ以外の国の人には削除される）。
- V3で何番がどの国かは、データ画面で、画面下の変数ビューをクリックして、V3の内容を見る。またはV3の度数分布を出す。
- このシンタックスのみを実行しても何も起きないが、この後に、クロス表シンタックスを書いて実行すると、人数が減っている。

11. データファイルの種類

- テキスト形式データファイルの場合 … まず、データ読み込み命令文 (DATA LIST FILE 文) を書く
シンタックス冒頭のデータの場所 (ドライブ名やフォルダ名) の指定文を書き換えてから、すべてのシンタックスを実行。
- SPSS形式データファイルがある場合 … SPSSでデータを開いてから、短い分析用シンタックスを実行するだけでよい。
シンタックスで GET FILE 文を書いて読み込むこともできる。
- テキスト形式データを使ってもよい。エクセル形式その他のデータもSPSSで簡単に読み込むことができる。

12. データの場所 シンタックスの書き方

- シンタックス冒頭にあるデータ読み込み命令文のデータの場所、例えばデータがHドライブに入っている場合、C:¥ を H:¥ に書き換える。

パソコンの多くは、ABがフロッピー、Cがハードディスク、D以降がDVD/CDドライブになっている。

- 例えば、データの場所が、CドライブのDADAフォルダの場合、シンタックスにC:¥DADA¥ と書き、その後、使うデータファイル名を正確に書く
フォルダを使っていない場合はC:¥ のみでよい

- Tanakaというユーザー名でMy Documentsフォルダにreidata.savというファイルが入っている場合

C:¥Documents and Settings¥tanaka¥My Documents¥reidata.sav

- Tanakaというユーザー名でデスクトップにreidata.savというファイルが入っている場合

C:¥Documents and Settings¥tanaka¥デスクトップ¥reidata.sav

- Hドライブのkamokuフォルダにreidata.savというファイルが入っている場合

H:¥kamoku¥reidata.sav

参考文献

- 村瀬洋一・高田洋・廣瀬毅士編. 2007. 『SPSSによる多変量解析』オーム社 .
SPSSの基礎の基礎からクラスター分析やログリニア分析まで解説。
- ボーンシュテット・ノーキ著 = 海野道郎・中村隆監訳 .1990 .
『社会統計学 - 社会調査のためのデータ分析入門』ハーベスト社 .
社会統計学に関する分かりやすい本
- 岸学. 2005. 『SPSSによるやさしい統計学』オーム社 .
- 酒井麻衣子 .2001 . 『SPSS完全活用法 データの入力と加工』東京図書 .
- 三宅一郎・山本嘉一郎他 .1986 . 新版SPSSX 基礎編』東洋経済新報社 .
- 海野道郎・原純輔 .2004 . 『社会調査演習 改訂版』東京大学出版会 .
社会調査法その他、クロス集計、エラボレイションなども解説あり

その他、村瀬研ホームページなどのシンタックス例文を参照。