

# SPSSによるクロス集計 操作のキソのキソ

村瀬 洋一

2021/4

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

## 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、村瀬ゼミホームページの、[資料その3](#)などを見てください(その他のデータファイルも可)。
- 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。拡張子が表示されない設定の場合は、設定を変えること。
- 保存先は自分のハードディスク内(立教ではホームディレクトリ、Hドライブと同じ)などにしてください。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。

2

## 2. SPSSの起動

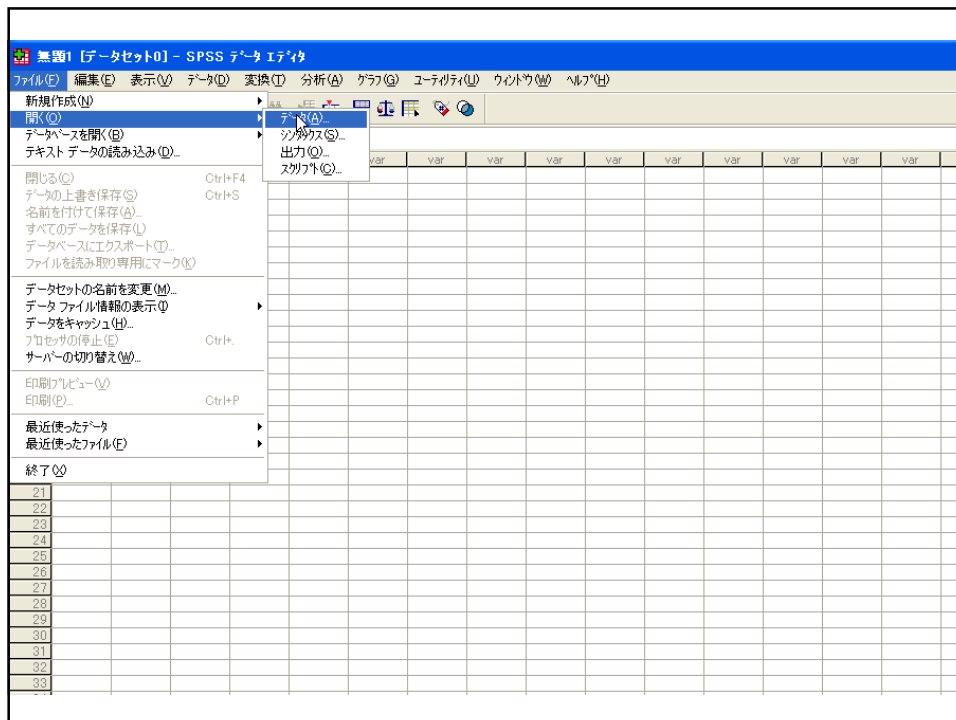
- スタートボタンを押してSPSSを起動し、画面の上にある「ファイル」をクリックして、先ほど保存したデータファイルを開く。  
なお、最初に出てくる「チュートリアル実行」画面はキャンセルボタンを押して消せばよい。
- データファイルをダブルクリックしてSPSSを起動してもよいが、エラー表示が隠れていて動かないことがある。その場合、隠れているエラーのOKボタンを押してから、SPSS画面上の「ファイル」をクリックして「開く」を選びデータファイルを開くこと。
- ホームページ上で「開く」をクリックしても動かないことがあるので、必ずスタートボタンを押してSPSSを起動してから、上記のように操作する。

3

## SPSSの画面

- データ画面
- シンタックス画面
- 出力画面(分析の操作をすると結果が出る)  
この3つの画面があることを理解すること。
- シンタックス画面には、自分で命令文を書く。ただし全角空白を、決して入れないこと。これがあるとエラーになって止まる。また、ピリオドを書く場所に注意する。
- まず、何らかのデータを開いてみる。

4



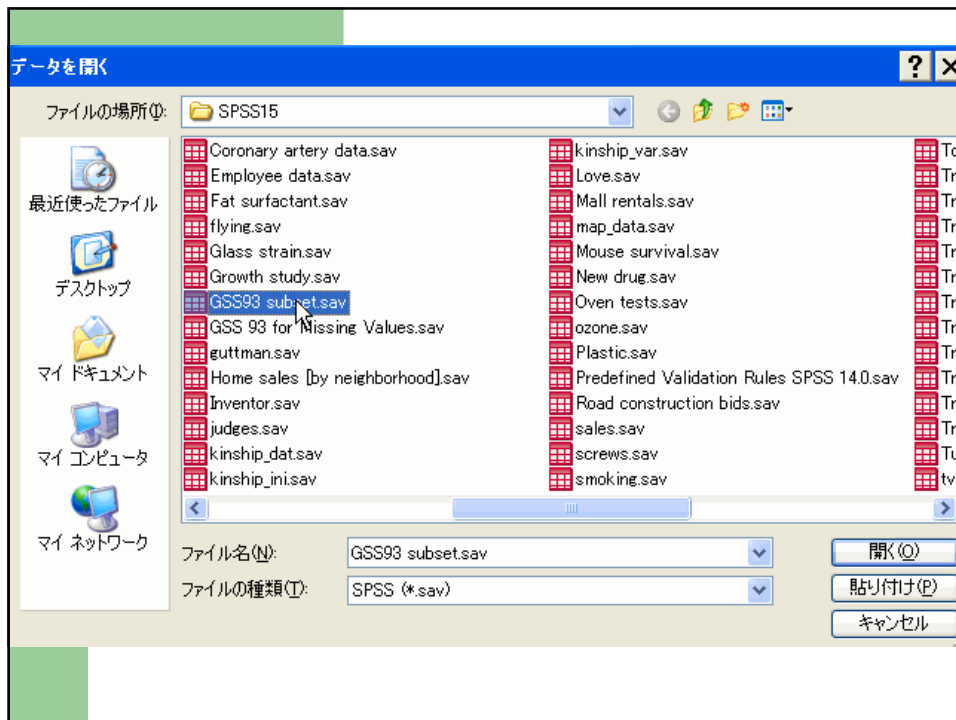
5

● まずは、「SPSSフォルダ」の中にあるGSS93subsetという練習用データなどを使う(次画面参照)。何らかの社会調査データなら良い。

● GSS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。

6



7

### 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。  
シンタックスを書くと、効率よく分析できる。
- クロス集計用のシンタックスを書く(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。

8

- 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も同じ結果になる。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。

9

この画面はSPSSの「分析」メニューから「クロス集計」を選択し、「セル」ボタンを押して「列%」(または行%)のみを出す設定をする様子を示しています。また、「セル」ダイアログボックスの「表示」タブで「列%」を選択し、「観測値」のチェックが入っている状態も確認できます。

回答番号	労働状況	婚姻状況	初婚年齢	最終学歴	父親学歴	母親学歴	性別	人種	総取
1	1	1	3	20	1	7	0	1	1
2	2	1	5	0	3	0	0	1	2
3	3	1	3	25	1	1	0	2	1
4	4	2	5	0	1	1	8	2	1
5	5	5	5	0	4	1	2	2	1
6	6	5	1	25	1	0	0	1	1
7	7	1	1	22	10	7	12	1	0
8	8	5	1	24	11	9	12	1	0
9	9	1	3	22	7	4	18	4	1
10	10	2	5	0	3	12	18	4	0
11	11	1	5	0	4	2	18	4	0
12	12	1	5	0	10	8	15	1	7
13	13	1	1	31	99	99	12	1	1
14	14	5	4	24	3	1	4	0	7
15	15	4	5	0	3	12	10	0	8
16	16	1	5	0	11	8	14	1	7
17	17	7	5	0	3	12	8	0	9
18	18	1	1	22	1	10	15	1	1
19	19	1	1	32	9	7	16	3	1
20	20	1	1	24	4	2	16	3	0
21	21	3	1	24	5	2	36	6	3
22	22	2	1	23	0	3	44	8	5
23	23	5	2	25	2	2	80	5	2
24	24	1	5	0	5	1	32	7	5
25	25	2	1	32	7	4	37	9	6
26	26	1	1	40	5	0	49	9	6
27	27	1	3	16	6	2	45	11	9
28	28	1	5	0	2	0	36	4	2
29	29	2	5	0	5	0	23	3	1
30	30	1	5	0	1	0	29	5	2
31	31	1	5	0	3	1	30	2	11
32	32	1	5	0	8	1	40	7	4
33	33	1	5	0	8	2	26	9	6
34	34	5	1	21	3	4	66	2	12
35	35	4	1	22	11	1	53	10	8
36	36	2	1	19	2	0	20	12	9
37	37	1	3	24	4	2	63	7	4
38	38	2	1	28	7	3	41	9	6
39	39	1	3	19	3	1	33	8	5
40	40	1	1	1	10	10	12	1	0
41	41	1	1	1	4	16	3	3	3
42	42	2	1	26	2	0	29	6	3
43	43	1	5	0	4	0	45	2	11
44	44	1	0	2	0	2	31	3	1
45	45	1	0	0	2	0	31	3	1

10

さらに、「層ボックス」に地区などを入れると、地区別に、学歴と性別のクロス集計表が出る

11

## 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。

12

## 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。

13

## 6. 分析の手順

- まず、自分が分析したい間(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間(説明変数Xとなるもの)について分析します。Xが基本属性でもいいでしょう。
- なお、必要に応じてカテゴリー合併を行うこと。

14

## 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。

15

## 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。

16



## 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。

17

## 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。

18

## 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。

19

## 12. データの場所(パス) 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は、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**DATA5フォルダ**の場合、シンタックスにC:¥DATA5¥ 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

20

- 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

21

## Mac OSの場合

- データファイルの場所(パス)の出し方を、別途調べると良い。

「マック ファイルの場所 パス」で検索

- MacのFinder上でパスを表示するには  
Finderを開いて「表示」→「パスバーを表示」
- ファイルパスをコピー  
option + command + Cキー

22

## 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。