

# SPSSシンタックスによる データ加工と分析入門

2009/10 立教大学社会学部  
村瀬 洋一

## 1. Syntaxとは何か

- データ加工や分析などの命令文(プログラム)  
クリックせずに、自分の操作を記述できる
- SPSSで、シンタックスのウィンドウを開く  
画面上「File」→新規作成→シンタックス

## シンタックスの長所

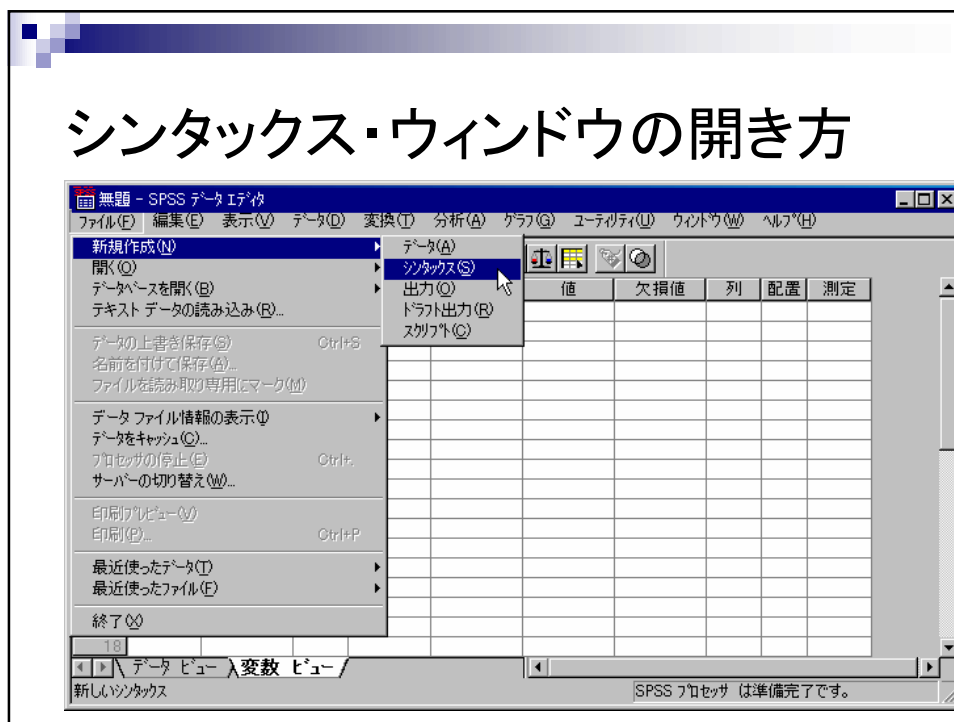
- 大量の操作(データ加工、分析)ができる
- 自分が実行した操作を保存できる
- 後で同じ操作を再現できる

## 2. 使い方

- シンタックス・ウィンドウを開く  
SPSSには、シンタックス・ウィンドウの他、データウィンドウ、出カウィンドウがある
- 自分でシンタックスを書く
- 実行する(実行ボタンを押す)



## シンタックス・ウィンドウの開き方



- この後、自分でシンタックスを書いた後に、実行する

実行ボタンを押すか、あるいは、

[ctrl+A\(全選択\)](#), [ctrl+R\(実行\)](#)

- シンタックスは、好きな名前を付けて保存する

## シンタックスの例

```
reil.sps - SPSS シンタックスの例
ファイル 編集 表示 データ 変換 分析 ケア エラーリポート 実行 ウィンドウ ヘルプ
DATA LIST FILE='C:\Documents and Settings\murase\My Documents\data2009\da09.txt' FIXED RECORDS=2
/1 ID 1-4 Q1 5 Q2 6 Q3 7 Q401 8 Q402 9
Q403 10 Q404 11 Q501 12 Q502 13 Q503 14 Q504 15 Q505 16 Q506 17
Q507 18 Q508 19 Q509 20 Q510 21 Q511 22 Q512 23 Q513 24 Q514 25
Q601 26 Q602 27 Q603 28 Q604 29 Q605 30 Q701 31 Q702 32 Q703 33
Q704 34 Q705 35 Q801 36 Q802 37 Q803 38 Q804 39 Q805 40 Q806 41
Q807 42 Q808 43 Q809 44 Q810 45 Q901 46 Q902 47 Q903 48
Q904 49 Q905 50 Q906 51 Q10 52 Q11 53 Q12 54 Q1301 55 Q1302 56
Q1303 57 Q1304 58 Q1305 59 Q1306 60 Q1307 61 Q1308 62 Q1309 63
Q1310 64 Q1311 65 Q1312 66 Q1313 67 Q1401 68
Q1402 69 Q1403 70 Q1404 71 Q1501 72 Q1502 73 Q1503 74 Q1504 75
/2 Q1505 1 Q1506 2 Q1507 3 Q1601 4 Q1602 5 Q1603 6
Q17 7 Q1801 8 Q1802 9 Q1803 10 Q19 11
Q2001 12 Q2002 13 Q2003 14 Q2004 15 Q2005 16 Q2006 17
Q21 18-19 Q22 20 Q2301 21 Q2302 22 Q2303 23
Q2304 24 Q2305 25 Q2306 26 Q2307 27 Q2308 28 Q2309 29
Q2310 30 Q2311 31 Q2312 32 Q2313 33 Q2314 34 Q2315 35
Q2316 36 Q2317 37 Q2318 38
Q2401 39 Q2402 40 Q2403 41 Q2404 42 Q2405 43 Q2406 44 Q25 45
Q2601 46 Q2602 47 Q2701 48-49 Q2702 50-51 Q2801 52 Q2802 53
Q30 54 Q31 55 Q32 56 Q33 57
Q3401 58 Q3402 59-60 Q35 61 Q36 62 Q37 63-64 Q38 65-66 Q39 67 Q2901 68-69 Q2902 70-71
Q2903 72-73 Q2904 74-75.

/***** NENREI TO NENREI KATEGORI SAKUSEI *****/
COMPUTE AGE =Q3402.
|
COMPUTE NENDAI =AGE.
RECODE NENDAI(19 THRU 29 =2)(30 THRU 39 =3)(40 THRU 49 =4)
(50 THRU 59 =5)(60 THRU 90 =6) .

Fre
VARIABLES = NENDAI.
```

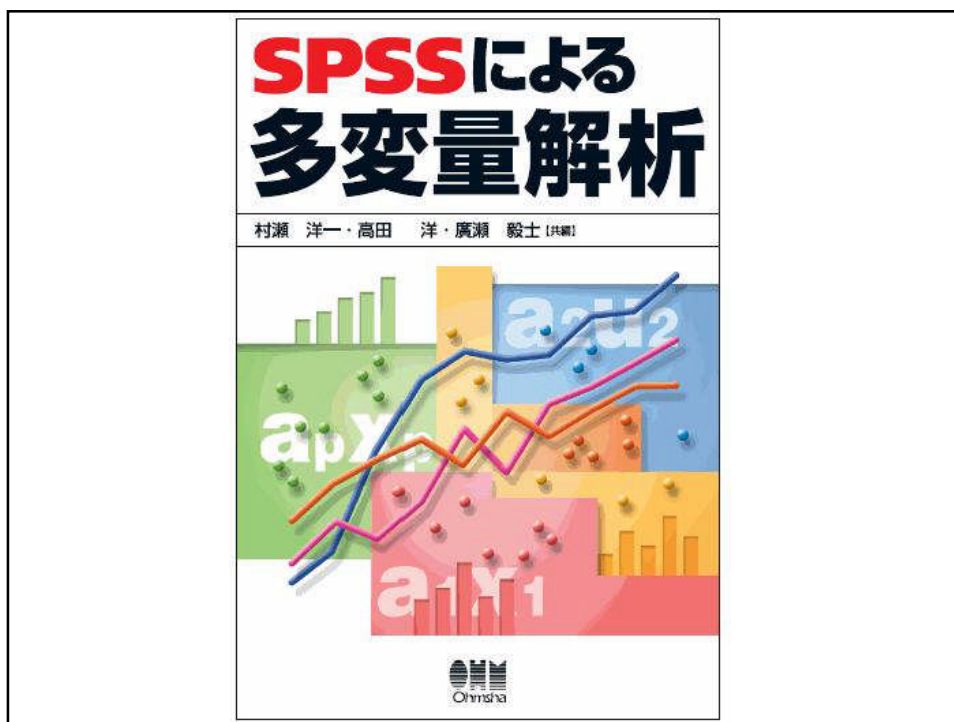
### ■ 書くときの注意点

- シンタックスは大文字と小文字を区別しない。
- 半角文字のみ。全角文字は使わない。
- 全角空白があるとエラーが出て止まってしまうので注意。
- 命令文の最後に1つだけピリオドをつける。

### ■ 変数名(調査ならば質問項目の名前)は、

実在する名前を正確に書くこと。変数名は  
SPSSのデータ・ウィンドウで確認すればよい。

詳しくは『SPSSによる多変量解析』オーム社 p.47参照



### 3.1. データ読み込み命令文 テキスト形式データファイルの場合

普通、ハードディスクはCドライブ  
マイドキュメント内のdata2009 というフォルダ内にある  
da09.txtファイルを読み込め、という命令文  
username はWindowsにログインした時のIDを入れる

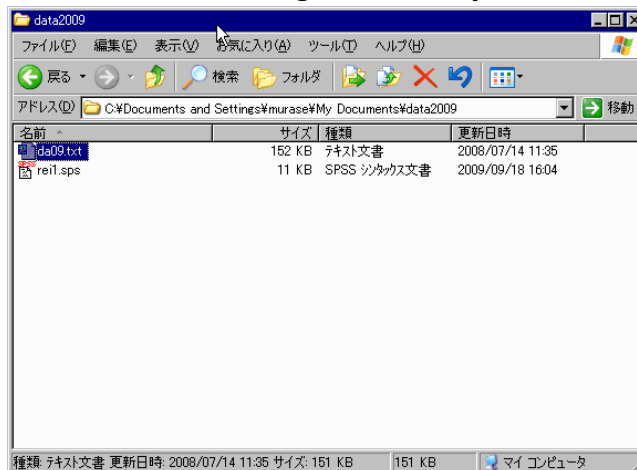
```
DATA LIST FILE=' C:¥Documents and Settings¥username¥  
My Documents¥data2009¥da09.txt' FIXED RECORDS=2  
/1 ID 1-4 Q1 5 Q2 6 Q3 7 Q401 8 Q402 9 ...
```

DATA LIST FILE シンタックスによりデータファイルを開く

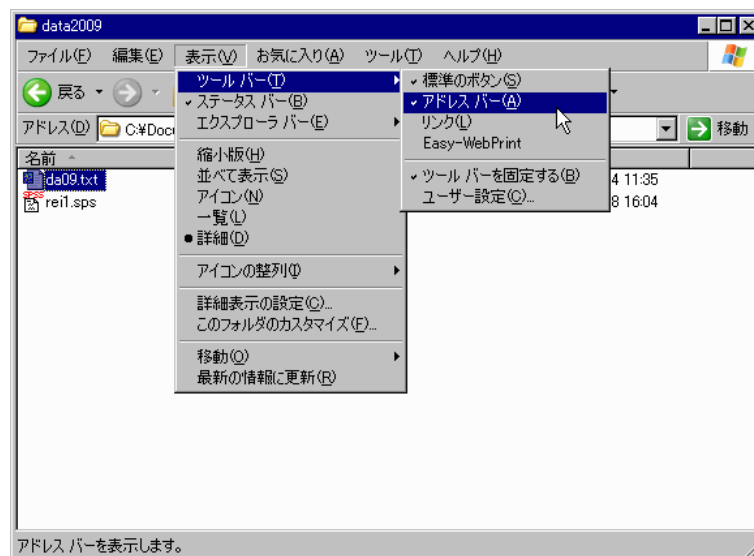
マイドキュメント内のdata2009というフォルダの中にあるda09.txtというファイル ... のアドレスは？

各フォルダのアドレスバーを見る

C:\Documents and Settings\murase\My Documents\data2009



## アドレスバーが出ない時



## 3.2.データ読み込み命令文 SPSS形式データファイルの場合

普通、ハードディスクはCドライブ

- マイドキュメント内のdata2009 というフォルダ内にある dreis.sav ファイルを読み込め、という命令文

```
GET FILE='C:\Documents and Settings\username\My Documents\data2009\dreis.sav'
```

- Hドライブ内のmydata というフォルダ内にある dreis.sav ファイルを読み込め、という命令文

```
GET FILE='H:\mydata\dreis.sav'
```

GET FILE シンタックスによりデータファイルを開く

## データ・ウィンドウの例

The screenshot shows the SPSS Data Editor window for 'dreis.sav'. The data grid contains 16 rows and 16 columns. The columns are labeled ID, Q1, Q2, Q3, Q401, Q402, Q403, Q404, Q501, Q502, Q503, Q504, Q505, Q506, Q507, and Q50. The data values are integers ranging from 1 to 4. The status bar at the bottom indicates 'SPSS プロセッサ は準備完了'.

	ID	Q1	Q2	Q3	Q401	Q402	Q403	Q404	Q501	Q502	Q503	Q504	Q505	Q506	Q507	Q50
1	101	1	3	4	1	1	1	3	4	3	3	2	2	3	2	
2	102	2	2	3	3	2	1	2	2	3	2	2	2	3	1	
3	103	2	2	2	2	2	2	2	3	2	1	1	1	2	1	
4	104	2	4	4	2	1	1	1	4	4	4	1	1	1	1	
5	105	2	3	2	1	2	1	1	3	3	2	2	3	1	1	
6	106	3	4	3	2	1	2	2	3	2	2	1	2	3	2	
7	107	2	2	3	4	2	1	2	3	4	1	2	2	3	2	
8	108	1	4	4	3	1	1	2	4	4	2	3	1	1	1	
9	109	2	2	3	2	1	2	2	2	2	3	3	2	2	1	
10	110	3	3	4	2	2	2	3	1	3	2	1	1	1	1	
11	111	2	3	3	2	2	1	2	3	2	2	3	2	2	2	
12	112	1	4	3	4	2	2	3	4	4	4	4	2	2	3	
13	113	2	2	2	4	1	2	1	3	2	2	1	1	1	1	
14	114	2	2	3	2	2	1	2	4	3	3	1	2	1	1	
15	116	2	2	4	4	3	2	2	2	3	3	3	1	9	3	
16	117	2	2	2	2	2	2	2	3	3	3	2	2	2	2	

## 4. データの中身について

- 横1行が1人の回答  
1000人の調査回答なら1000行ある  
Q1の質問項目(変数)で1と答えたなど
- 4段階回答だが、9と答えている人がいる  
非該当8、無回答(DK/NA)9 とすることが多いため  
DK/NA: Don't Know, No Answer
- 複数回答  
いくつでも○をつけてください、という問の場合  
○をつけた人は1, なければ0, という1,0のデータになっている

## 5. データ加工の例

### リコード文によるカテゴリ合併

- 例 変数Q2の4段階回答を2段階にする。  
以下は、Q2という変数をもとに、N2という新しい変数を作成。  
新変数名がN2ということを理解する。  
リコード文によりN2の回答を2段階にする例。

```
COMPUTE  N2 = Q2 .  
RECODE  N2 (1, 2=1) (3, 4=2) .
```

```
CRO                                     ← クロス集計命令
```

```
  /TAB = N2 by Q2
```

```
  /CEL = ROW .
```

- 上記のシンタックスを実行すると、N2を用いたクロス集計結果が出る。変数名は、自分の使いたい変数の名前を書けばよい。
- リコード文のみを実行しても何もおきないように見えるが、その後に新変数を使って分析を行えばよい。



## 例 収入カテゴリーの作成

- 収入の問がQ15 だとする  
回答は1～14までの選択肢

```
COMPUTE ZAISANCT=Q15 .  
RECODE ZAISANCT (1 thru 4 =1) (5, 6, 7=2)  
(8, 9=3) (10, 11=4) (12, 13, 14=5) .
```

## 6. 例 年齢を10歳刻みに分けるには

- リコード文によるカテゴリー合併の例  
AGEという細かい変数を、NENDAIという5段階の変数（新しい変数）に変換。  
新変数名がNENDAIということを理解する。

```
COMPUTE NENDAI =AGE.  
RECODE NENDAI (1 THRU 29 =2) (30 THRU 39 =3)  
(40 THRU 49 =4) (50 THRU 59 =5) (60 THRU 97 =6) .
```

```
CRO  
/TAB = NENDAI by Q3  
/CEL = ROW .
```

- AGEの部分は、年齢の変数名に書き変える
- Q3 の部分は、自分が分析したい変数名に書き変える  
上記のシンタックスを実行すると、年代ごとの集計が出る。

## 7. 例 学歴を教育年数に変換するには

- リコード文による値の書き換えの例

Q16という本人学歴の変数を、EDUという教育年数の変数（新変数）に変換。

以下のシンタックスを書いて実行すればよい。

```
COMPUTE    EDU=Q16.  
RECODE    EDU (1=6) (2=9) (3=12) (4=13)  
           (5=14) (6=16) (7, 9=99).  
MISSING VALUES EDU (99).
```

1と答えた人（小学校卒）なら6年、などに変換している。  
7のその他や9の無回答は欠損値としている。

- Q16の部分は、学歴の変数名に書き変える。
- これだけでは何もおきないが、この後に新変数EDUを使って、クロス集計など分析を行えばよい。

- 無回答は普通、9や99を入力している。
- そのままだと、分析上問題が起きるので、上記のようにMISSING VALUES コマンドを使い、欠損値処理をする。
- きちんと新変数ができたかどうか、単純集計やクロス集計表を出して、新変数の内容を確認してから分析すること。

## 8. 男女別にデータを分割

```
SORT CASES BY Q21 .  
SPLIT FILE BY Q21 .
```

- Q21の部分は、性別の変数名に書き変える
- データ・ウィンドウを見ると、全員のデータが性別順に並べ替えられている。
- これを実行した後に分析を行うと、すべての結果が男女別に出る。

## 9. ある値の人にデータを限定

- SELECT文を使うと、ある条件の人のみのデータとなる。
- データ人数が減るので注意。
- 例 70歳未満の人のみを使って分析したい場合

```
SELECT IF AGE < 70.
```

AGE 部分は存在する変数名を正確に書く

## 10. IF文の使い方

- IF文を使うと、ある条件の場合のみ、何らかの命令を実行可

- 例 年齢が30歳以上の場合に、変数Q7Aの値に1を足す  
( )内に条件を書き、その後に、条件に合う場合に  
実行したい命令を書けばよい

```
IF (age >29) Q7A=Q7A+1.
```

- 例 変数Q25の値が2か3の場合に、  
変数SENMONの中身を1にする。

```
COMPUTE SENMON =0.
```

```
IF (Q25=2) SENMON =1.
```

```
IF (Q25=3) SENMON =1.
```

Q25が2か3の場合のみ、 SENMONの中身は1、あとは0になる。

## 11. 分析命令文の見本

- 単純集計シンタックス見本

```
FRE VAR = Q2 Q3 .
```

- クロス集計シンタックス見本

```
CRO
```

```
/TAB = V3 by V29
```

```
/CEL = ROW
```

```
/STA .
```

/CEL の行に COL を書くと縦%、COUを書くと実数が出る。

■ 重回帰分析の例

被説明変数がQ6Bの場合

```
REG /DEP Q6B  
/METHOD=ENT Q2 Q3 Q4A .
```

■ 分散分析の例

被説明変数がEDUの場合

```
UNI EDU by sex nendai  
/PLOT = PRO( nendai * sex ) .
```

- 変数名は自分が分析したい変数に書き換える。
- 最後に1つだけピリオドがつくのを忘れずに。

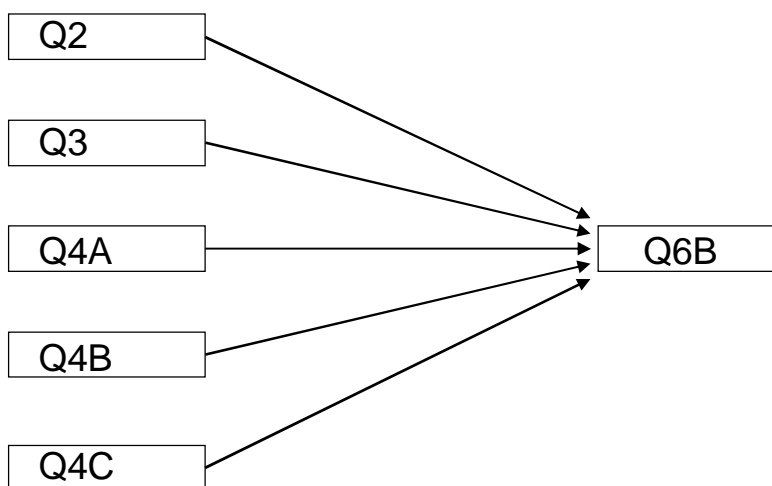


図1 重回帰分析のモデル

## 重回帰分析の注意点

- 多重共線性に注意する  
X同士の相関を確認する
- 分析前に
  - 変数の方向をそろえる
  - 欠損値処理をする  
まず、用いる変数の分布を確認すること
- なお、男女別、国別にデータを分割後に分析した方が、説明力が上がることが多い。
- YもXがすべて連続量という点が、分散分析と異なる。できるだけ細かい変数を使うこと。

## 12. 変数や値に名前をつける

- 変数に名前をつける  
VARIABLE LABELS  
Q1 '年齢'  
/ Q2 'Life Satisfaction'
- 値に名前をつける  
VALUE LABELS Q6A TO Q6E  
1 'agree'  
2 'somewhat agree'  
3 'somewhat disagree'  
4 'disagree'  
9 '無回答'

その他、詳しくは「応用調査実習」ホームページ参照