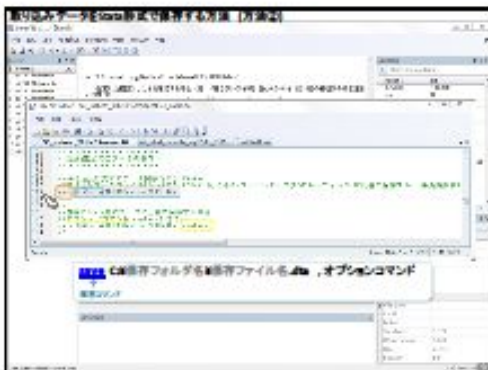


取り込んだデータの保存

ExcelまたはCSVファイルからStataに取り込んだデータをStata形式で保存する。

- ▶ Stata形式ファイル (.dta)での保存



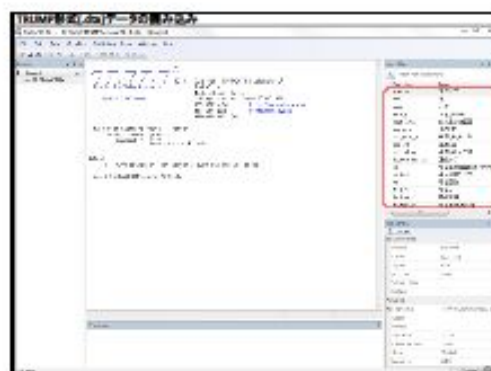
TRUMP形式(.dta)データの読み込み

Stata形式(.dta)のデータを開く

- ▶ フォルダ内に保存されたファイル(.dta)を直接クリックして開く
- ▶ ファイル(.dta)の読み込みコマンドにて開く

TRUMP形式(.dta)データの読み込み

The diagram shows a file explorer window with a folder icon. Inside the folder, there is a .dta file icon. A red box highlights the .dta file icon, and a callout bubble with the text 'ダブルクリック' (Double-click) points to it.



ログファイルの作成

ログファイルとは…
 Stataによる作業結果を記録したファイル

コマンド実行などの作業結果はResult windowに表示されるが、いくつかの作業を行うと結果が長くなり、すべて選んで表示されない。(最初の結果から消滅されている)

ログファイルを作成すると、
 ▶すべての結果を保存可能
 ▶text形式で保存し、Excelファイルから開くと、結果tableを作成するのに便利

ログファイルの作成: Do-file editor

log using logファイル名
 この間の作業結果を記録する

ログファイルの作成: Do-file editor

log using ファイル名
 ファイル名
 ファイル名

ログファイルの作成: result window

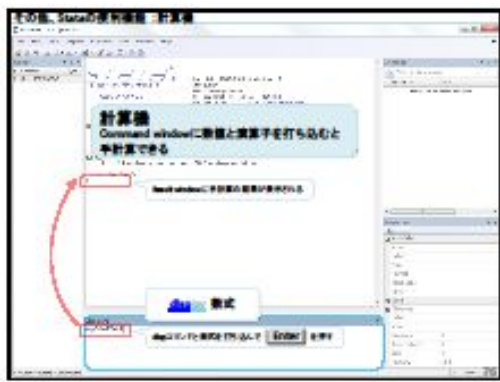
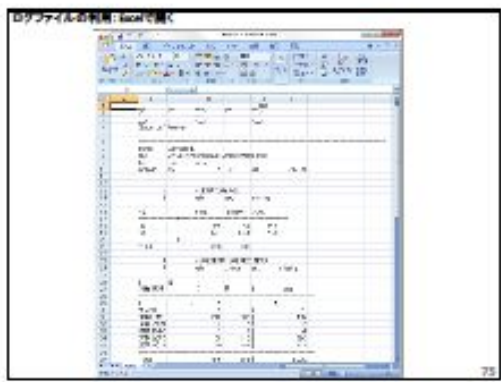
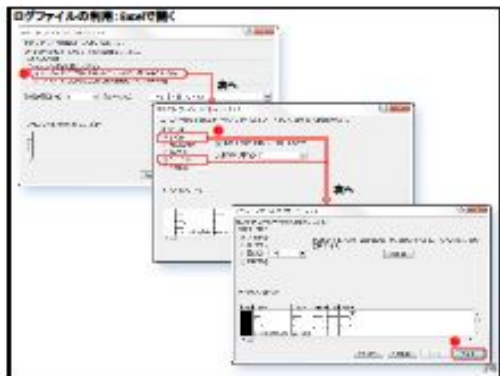
log using ファイル名
 ファイル名
 ファイル名

ログファイルの保存

デフォルトでは、使用データファイルの同じフォルダ内に、ログファイルは保存されます。(標準的フォルダ構成のため)

ログファイルの利用

ログファイルを作成すると、
 ▶すべての結果を保存可能
 ▶text形式で保存し、Excelファイルから開くと、結果tableを作成するのに便利



* データを取り込む前の準備 *

///
* データセットの前準備 *///

- * Excel または CSV のファイル名を英語名にしてください。(TRUMP2013 など)
- * 項目が英語名のデータセットを使用してください。
- * 2010 年度データセットに関しては、JSHCT HP に英語項目名変換用のエクセルマクロを使用してください。(JSHCT ホームページよりダウンロード可能)
- * 最終行に入力率が入っている場合はこれを削除してください。
- * WG データセット配布時に設定されている開封パスワードを解除し、保存したデータを使用してください。

* データの取り込み *

///
* Excel ファイルからの取り込み *///

- * Stata12 からはエクセルファイルのままで取り込みが可能です。
- * Excel のファイルサイズが 40 メガバイトを超える場合には、Excel ファイルからの取り込みはできません。

///
* CSV ファイルからの取り込み *///

- * 横断的 WG でフルデータを取り込む場合など、Excel ファイルの大きさが 40 メガバイトを超える場合は Excel ファイルからの取り込みができません。
- * CSV ファイルから取り込んでください。

*****データ取り込みコマンド*****

* insheet using C:¥データファイルの保存先フォルダ名¥データファイル名, オプションコマンド

***** 以下に記載するコマンドの解説 *****

- * 取り込みコマンドの using 以降はファイルの場所(directory)を示します。
- * directory は、あらかじめローカルディスク C の中に"data"というフォルダを作成し、元データを保存してある場所です。
- * オプションコマンドの name は、第一行のデータが項目名であることを示す。

* ちなみに、オプションコマンドとは、カンマ以降のコマンドのことです。

***** ミニ知識 *****

* 米印が先頭に来る場合、あるいは/*と*/で囲んだところは、メモ書きとしてStataは認識しません。

***** CSV ファイルからの取り込みコマンド *****

*for Windows

```
insheet using C:\data\TRUMPSeminar2013.csv, name
```

*for Mac (OS の version によって取り込み方法が異なる場合があります)

```
insheet using /c/data/TRUMPSeminar2013.csv, name
```

*読み込みに多少時間はかかります、あせらないでくださいね！！

***** Excel ファイル(xls ファイル)からの取り込みコマンド *****

```
import excel using C:\data\TRUMPSeminar2013.xls, firstrow
```

***** エクセルファイル(xlsx ファイル)からの取り込みコマンド *****

```
import excel using C:\data\TRUMPSeminar2013.xlsx, firstrow
```

* 変数にラベルを付ける *

*取り込んだ変数の内容がわかるように変数ラベルを付けると便利です。

///* ラベルの付け方の一例 *///

/*変数 survival に日本語ラベル"生存状況"を付けます。

```
label var survival "生存状況"
```

*変数 sex に日本語ラベル"性"を付けます。

```
label var sex "性"
```

*変数 race に日本語ラベル"人種"を付けます。

```
label var race "人種"
```

*/

*JSHCT TRUMP データを用いた登録研究の解析を行う場合、各 WG のデータセットに対応した日本語ラベルスクリプトをホームページよりダウンロードできます

***** do ファイルを使って、ラベルを付けるコマンドを実行する *****

*配布した Do ファイル[01_label_variable_wg17-21_XXXXXX.do]を Do-file Editor にて開き、実行してください。

* 取り込んだデータの確認 *

***** 取り込んだデータの確認方法 *****

* 取り込んだ内容を直接確認する

* Data Editor(Browse)を開いて全内容を見ることができます

* 標本数と変数の属性を確認する (各変数のタイプ、ラベル名、観測値、変数の数、ファイルサイズなど)

*使用するコマンド : describe を用いる (例 : 変数 survival ,no, tx_type)

de survival no tx_type

* 連続変数の要約値を確認する (指定した連続変数の観測値、平均値、標準偏差、最小値/最大値)

*使用するコマンド : summarize (例 : 連続変数 age_month)

sum age_month

*オプションコマンド detail を加えると、中央値、パーセンタイル値、歪度、尖度を算出することができます。

sum age_month, detail

* カテゴリー変数の要約値を確認する (各カテゴリーの頻度、割合)

*使用するコマンド : tabulate (例 : カテゴリー変数 tx_type)

tab tx_type

*オプションコマンド missing を加えると、欠損値のデータ数を算出することができます。

tab tx_type, missing

*tab コマンドの後ろに 2 つのカテゴリー変数を指定するとクロス集計を行うことができます。

tab tx_type no

* データのサンプル数(観測値)を確認する

*使用するコマンド : count

```
count
```

*リストを作成してデータの内容を確認する

*使用するコマンド : list

*指定した範囲のリストを確認する場合 : if 条件式

** (例:移植日が2000年、2001年の移植例の移植種類と移植日)

```
list tx_type tx_date if tx_date>="2000/1/1" & tx_date<="2001/12/31"
```

```
/*
```

result window に表示される内容が多い場合、result window に -more- の文字が表示されます。

-more- の文字上をクリックするか、Enter キーまたは Space キーを押すと結果の続きが表示されま
す。

```
*/
```

* Stata 形式でのデータの保存 *

*****Stata 形式でのデータ保存コマンド*****

**取り込みデータファイル(Excel または CSV)と同じ場所(ローカルディスク C の data フォルダ)に同名
で保存する。(半角英数字)

```
save C:\data\TRUMPSeminar2013.dta
```

```
/*
```

**既存の Stata 形式データに上書き保存する場合

**オプションコマンド replace を付ける

```
save C:\data\TRUMPSeminar2013.dta, replace
```

```
*/
```

* Stata 形式データの読み込み *

***** Stata 形式で保存した .dta ファイルは以下の方法で読み込む *****

* フォルダ内に保存されたファイル(.dta)を直接クリックして開く

* .dta ファイルの読み込みコマンドにて開く

```
*use C:\data\TRUMPSeminar2013.dta, clear
```

```
/* オプションコマンドの clear を付けると、既に読み込んであるデータを Stata のメモリから消して、同時に新たなデータを読み込むことができます。
```

```
(開いている Stata を閉じることなく、新たな保存データ.dta を読み込む)
```

```
*/
```

* ログファイルの作成と利用 *

*ログファイルとは、Stata の作業結果を記録したファイルのことです。

```
///* ログファイルの作成 *///
```

*作業結果は Result window に表示されますが、いくつかの作業を行い結果が長くなると遡って表示されません。

*ログファイルを作成すると、すべての作業結果を保存することができ便利です。

```
///* ログファイルの利用 *///
```

*ログを text 形式で保存すると、メモ帳などで開くこともできます。

*ログを text 形式で保存し、Excel ファイルから開くと、結果 table を作成する際に便利です。

***** ログファイルの作成方法 *****

```
/*
```

```
log using ログファイル名
```

```
xxx
```

```
log off
```

```
yyy
```

```
log on
```

```
zzz
```

```
log close
```

```
translate ログファイル名.smcl ログファイル名.txt
```

```
*/
```

*ログファイルに「xxx」「zzz」の部分の記録はされますが、「yyy」部分は記録されません。

*ログファイルを作成し、Excel ファイルで table を作成

*例: データの要約値をログファイル[result]に記録して、table を作成する。

log using table

*性別の移植件数

tab sex, missing

*移植種別移植件数(性別)

tab tx_type sex, missing

*患者性別、年齢 (Mann-Whitney 検定)

ranksum age, by(sex)

*患者性別、移植種別 (chi-square 検定)

tab sex do_sex, chi2

log close

translate table.smcl table.txt

***** text ファイルで保存された log ファイルを Excel で開く方法 *****

/* Excel 開く text ファイル開く [カンマやタブなどの区切り文字によってフィールドごとに区切られたデータ]を選択

次へ [タブ] [スペース]にチェックを入れる 次へ 完了

*/

変数の作成の基本的事項

平成25年度「UMM教育セミナー」開催編
2013年10月20日

変数作成の基本的事項

名古屋大学附属病院
丸瀬町医療・臨床研究支援センター
松原 八千代

配布資料

- ① Minidataset_2013 (Stata Dataset)
- ② Minidataset_2013 (Excel)
- ③ SL_minidata_2013 (Stata Do-file + プリント)
- ④ ミニデータワークシート 2013 (Excel + プリント)



このセクションの目標

- 50個、少数の項目のデータ
- Stataを使用した簡単なデータクリーニング
- 変数作成
- 多変量解析前のデータ準備

minidataset_2013.xls



図形ファイルは、直接ドキュメントへアクセスして「マイクログラフ」フォルダ内に格納してください。

このセクションのゴール



- ✓ Table 1
患者数表の作成
- ✓ Figure 1
US curve の作成

minidataset_2013.dta



Stata関数とエクセルデータの取り込み

- Stataを起動
 - ログファイル実行式
 - 実行ファイルを開く
 - minidataset_2013.xlsをStataに取り込み
 - 実行ファイルを開く
- 参考: UMM研究開発センター 公開編訳集
SL_Minidata_2013 do file 7-9頁目

ログファイルの作成 ①

配布資料: S14Minidata_2013: 8行目



エクセルデータ取り込み (コマンド使用)

S14Minidata_2013: 7行目



新たな変数作成の練習 (洋字運動)

S14Minidata_2013: 11~28行目

`generate, replace`
 新たな変数の作成、データの置換

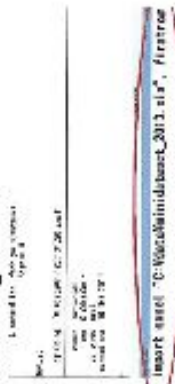
`drop, keep`
 変数・変例の削除

ログファイルの作成 ②



エクセルデータ取り込み (コマンド使用)

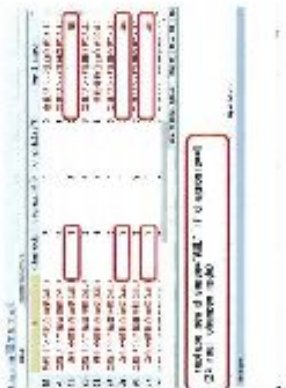
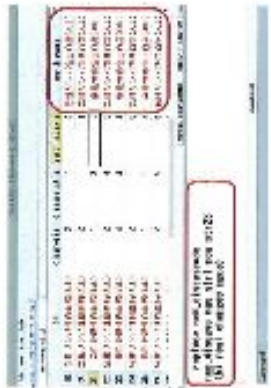
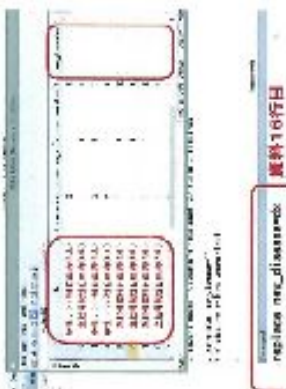
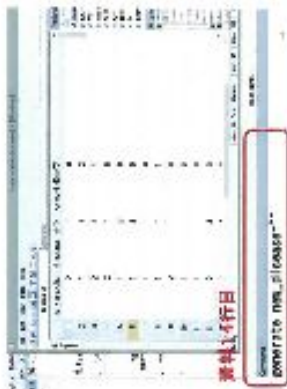
S14Minidata_2013: 7行目

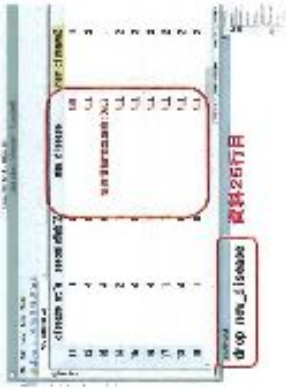
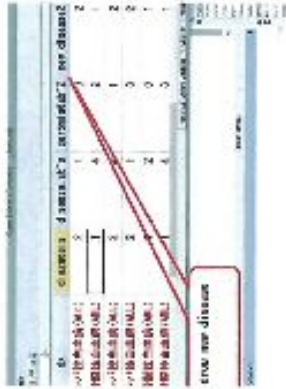
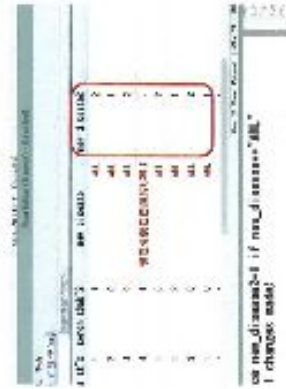


新たな変数作成の練習

S14Minidata_2013: 11~22行目







新たな変数作成の練習

実行結果: S3.Mimido_2013...23*28行目

```

23 /new_diagnoseの練習
24 /new_diagnoseの練習
25 drop raw_diagnose;
26 * new_diagnoseの練習
27 drop if diagnosed=1;
28 keep if diagnosed=1;

```


順番に疾患データの中央値を確認
58. tab diagnose, missing (欠損値も表に含める)

```
tab diagnose, missing
```

疾患名	Freq.	Percent	Cum.
1	54	46.11	46.11
2	53	45.25	91.36
3	3	2.64	94.00
Total	110	100.00	

labelが付いた状態

順番に疾患データの中央値を確認
59. tabdiag, missing (欠損値も表に含めない)

```
tabdiag, missing
```

疾患名	Freq.	Percent	Cum.
1	54	46.11	46.11
2	53	45.25	91.36
3	3	2.64	94.00
Total	110	100.00	


```
tab diagnose, missing
```

疾患名	Freq.	Percent	Cum.
1	54	42.11	42.11
2	53	42.43	84.54
3	3	2.46	87.00
Total	110	100.00	

2x2 tableを作る
61. tabdiag, missing

```
tabdiag, missing
```

疾患名	Freq.	Percent	Cum.
1	54	42.11	42.11
2	53	42.43	84.54
3	3	2.46	87.00
Total	110	100.00	

2x2 tableを作る
61. tabdiag, missing

```
tabdiag, missing
```

疾患名	Freq.	Percent	Cum.
1	54	42.11	42.11
2	53	42.43	84.54
3	3	2.46	87.00
Total	110	100.00	

S1XMiniData_2013 62行目～

```

*totalを100%にする*
*列下の列出し*
*変数間の関係は、**関係は、変数間の関係は**
* % and %*
| for 変数1
== (equal 変数1)
=> total 変数1 (変数1)
* keep if diagnose=1 | diagnose=2
* use if diagnose=1
* use if diagnose=2
* use if diagnose=1 & 2
* use if diagnose=1 & 2

```

66. keep if diagnose=1 | diagnose=2
68. count

```
keep if diagnose=1 | diagnose=2
```

疾患名	Frequency	Percent	Cum.
1	54	42.11	42.11
2	53	42.43	84.54
3	3	2.46	87.00
Total	110	100.00	

S1XMiniData_2013 70行目～

```

survival生存状況のデータと、
last_date最終観測日(欠損の値も併列)
変数は、*一部の観察
71. tab survival, missing
72. tab last_date, missing
*列に、73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000

```

71. tab survival, missing

```
tab survival, missing
```

変数名	Freq.	Percent	Cum.
1	54	42.11	42.11
2	53	42.43	84.54
3	3	2.46	87.00
Total	110	100.00	

71. tab last_date, missing

```
tab last_date, missing
```

変数名	Freq.	Percent	Cum.
1	54	42.11	42.11
2	53	42.43	84.54
3	3	2.46	87.00
Total	110	100.00	

73. drop if survival==0 | last_data=**
75. count

518MInData_2013 77行目

• 性別データの作成
 1. 性別データの作成
 2. 性別データの作成
 3. 性別データの作成
 4. 性別データの作成
 5. 性別データの作成
 6. 性別データの作成
 7. 性別データの作成
 8. 性別データの作成
 9. 性別データの作成
 10. 性別データの作成

性別データの確認
89. tab sex, missing

sex	Frequency	Percentage	Missing
Female	12	60.0%	
Male	8	40.0%	
Total	20	100.0%	

新たな変数"pc_sex"を作成。
 sex==1 女の性別は1
 90. gen pc_sex = 0 if sex==2

変数"pc_sex"の平均値を、
 sex==1 男の性別は1に修正する
 92. replace pc_sex = 1 if sex==1

変数"pc_sex"にラベル"pc_sex_label"を
 関連付ける(各社をつける作業)
 100. label values pc_sex sex_label

2x2 tableを作って、関連していないか確認
 107. tab sex pc_sex, missing

pc_sex	Female	Male	Total
sex	12	8	20
Total	12	8	20

518MInData_2013 106行目 - エラスティックネット

「sex_label」という名前のラベルの定義をする
 99. label define sex_label 0 "Female" 1 "Male"

166 tab year_cat = year_cat * \$!\$! * 1000000
 167 tab year_cat = year_cat * \$!\$! * 1000000
 168 tab year_cat = year_cat * \$!\$! * 1000000
 169 tab year_cat = year_cat * \$!\$! * 1000000

Year	Match	Mismatch	Total
2000	1	1	2
2001	1	1	2
2002	1	1	2
2003	1	1	2
2004	1	1	2
2005	1	1	2
2006	1	1	2
2007	1	1	2
2008	1	1	2
2009	1	1	2
2010	1	1	2
2011	1	1	2
2012	1	1	2
2013	1	1	2
2014	1	1	2
2015	1	1	2
2016	1	1	2
2017	1	1	2
2018	1	1	2
2019	1	1	2
2020	1	1	2
2021	1	1	2
2022	1	1	2
2023	1	1	2
2024	1	1	2
2025	1	1	2
2026	1	1	2
2027	1	1	2
2028	1	1	2
2029	1	1	2
2030	1	1	2

Table 1. の作成

Year	Match	Mismatch	Total
2000	1	1	2
2001	1	1	2
2002	1	1	2
2003	1	1	2
2004	1	1	2
2005	1	1	2
2006	1	1	2
2007	1	1	2
2008	1	1	2
2009	1	1	2
2010	1	1	2
2011	1	1	2
2012	1	1	2
2013	1	1	2
2014	1	1	2
2015	1	1	2
2016	1	1	2
2017	1	1	2
2018	1	1	2
2019	1	1	2
2020	1	1	2
2021	1	1	2
2022	1	1	2
2023	1	1	2
2024	1	1	2
2025	1	1	2
2026	1	1	2
2027	1	1	2
2028	1	1	2
2029	1	1	2
2030	1	1	2

170 sum age = sum age * 1000000
 171 sum age = sum age * 1000000
 172 sum age = sum age * 1000000

Age	Match	Mismatch	Total
18	1	1	2
19	1	1	2
20	1	1	2
21	1	1	2
22	1	1	2
23	1	1	2
24	1	1	2
25	1	1	2
26	1	1	2
27	1	1	2
28	1	1	2
29	1	1	2
30	1	1	2
31	1	1	2
32	1	1	2
33	1	1	2
34	1	1	2
35	1	1	2
36	1	1	2
37	1	1	2
38	1	1	2
39	1	1	2
40	1	1	2
41	1	1	2
42	1	1	2
43	1	1	2
44	1	1	2
45	1	1	2
46	1	1	2
47	1	1	2
48	1	1	2
49	1	1	2
50	1	1	2
51	1	1	2
52	1	1	2
53	1	1	2
54	1	1	2
55	1	1	2
56	1	1	2
57	1	1	2
58	1	1	2
59	1	1	2
60	1	1	2
61	1	1	2
62	1	1	2
63	1	1	2
64	1	1	2
65	1	1	2
66	1	1	2
67	1	1	2
68	1	1	2
69	1	1	2
70	1	1	2
71	1	1	2
72	1	1	2
73	1	1	2
74	1	1	2
75	1	1	2
76	1	1	2
77	1	1	2
78	1	1	2
79	1	1	2
80	1	1	2
81	1	1	2
82	1	1	2
83	1	1	2
84	1	1	2
85	1	1	2
86	1	1	2
87	1	1	2
88	1	1	2
89	1	1	2
90	1	1	2
91	1	1	2
92	1	1	2
93	1	1	2
94	1	1	2
95	1	1	2
96	1	1	2
97	1	1	2
98	1	1	2
99	1	1	2
100	1	1	2

173 sum age = sum age * 1000000
 174 sum age = sum age * 1000000
 175 sum age = sum age * 1000000

Age	Match	Mismatch	Total
18	1	1	2
19	1	1	2
20	1	1	2
21	1	1	2
22	1	1	2
23	1	1	2
24	1	1	2
25	1	1	2
26	1	1	2
27	1	1	2
28	1	1	2
29	1	1	2
30	1	1	2
31	1	1	2
32	1	1	2
33	1	1	2
34	1	1	2
35	1	1	2
36	1	1	2
37	1	1	2
38	1	1	2
39	1	1	2
40	1	1	2
41	1	1	2
42	1	1	2
43	1	1	2
44	1	1	2
45	1	1	2
46	1	1	2
47	1	1	2
48	1	1	2
49	1	1	2
50	1	1	2
51	1	1	2
52	1	1	2
53	1	1	2
54	1	1	2
55	1	1	2
56	1	1	2
57	1	1	2
58	1	1	2
59	1	1	2
60	1	1	2
61	1	1	2
62	1	1	2
63	1	1	2
64	1	1	2
65	1	1	2
66	1	1	2
67	1	1	2
68	1	1	2
69	1	1	2
70	1	1	2
71	1	1	2
72	1	1	2
73	1	1	2
74	1	1	2
75	1	1	2
76	1	1	2
77	1	1	2
78	1	1	2
79	1	1	2
80	1	1	2
81	1	1	2
82	1	1	2
83	1	1	2
84	1	1	2
85	1	1	2
86	1	1	2
87	1	1	2
88	1	1	2
89	1	1	2
90	1	1	2
91	1	1	2
92	1	1	2
93	1	1	2
94	1	1	2
95	1	1	2
96	1	1	2
97	1	1	2
98	1	1	2
99	1	1	2
100	1	1	2

181 tab year_cat = year_cat * 1000000
 182 tab year_cat = year_cat * 1000000
 183 tab year_cat = year_cat * 1000000

Year	Match	Mismatch	Total
2000	1	1	2
2001	1	1	2
2002	1	1	2
2003	1	1	2
2004	1	1	2
2005	1	1	2
2006	1	1	2
2007	1	1	2
2008	1	1	2
2009	1	1	2
2010	1	1	2
2011	1	1	2
2012	1	1	2
2013	1	1	2
2014	1	1	2
2015	1	1	2
2016	1	1	2
2017	1	1	2
2018	1	1	2
2019	1	1	2
2020	1	1	2
2021	1	1	2
2022	1	1	2
2023	1	1	2
2024	1	1	2
2025	1	1	2
2026	1	1	2
2027	1	1	2
2028	1	1	2
2029	1	1	2
2030	1	1	2

184 tab year_cat = year_cat * 1000000
 185 tab year_cat = year_cat * 1000000
 186 tab year_cat = year_cat * 1000000

Year	Match	Mismatch	Total
2000	1	1	2
2001	1	1	2
2002	1	1	2
2003	1	1	2
2004	1	1	2
2005	1	1	2
2006	1	1	2
2007	1	1	2
2008	1	1	2
2009	1	1	2
2010	1	1	2
2011	1	1	2
2012	1	1	2
2013	1	1	2
2014	1	1	2
2015	1	1	2
2016	1	1	2
2017	1	1	2
2018	1	1	2
2019	1	1	2
2020	1	1	2
2021	1	1	2
2022	1	1	2
2023	1	1	2
2024	1	1	2
2025	1	1	2
2026	1	1	2
2027	1	1	2
2028	1	1	2
2029	1	1	2
2030	1	1	2

177 tab mismatch = 1000000
 178 tab mismatch = 1000000
 179 tab mismatch = 1000000

mismatch	True	Percent	Sum
match	25	50.00	25.00
mismatch	25	50.00	25.00
Total	50	100.00	50.00

Figure 1. の作成

