

「EZRを使ってみよう 生物統計ハンズオンWebinar初級」

-事前学習資料-

Webinarの目的

**フリーの統計解析ソフト、EZRの基本的な
使用方法を習得すること**

1. 事前準備のお願い

事前準備のお願い

EZRのインストールと事前学習

EZRのインストールと事前学習

- EZRのインストール**未実施**の方

- 本資料の5ページに進んでください。 **Webinar当日まで**に5ページ以降をすべて実施いただき、EZRのインストールを完了、インストールしたEZRの動作確認をお願いいたします。
- Mac User は12ページ以降もご参照ください。

- EZRのインストール**実施済み**の方

- 本資料の15ページに進んでください。 **Webinar当日まで**にすべて実施いただくようお願いいたします。15ページ以降では、データファイルの作成と取り込み方法を紹介しており、インストールしたEZRの動作確認のために必要な作業です。

2.1 EZRのインストール

EZRについて

- **無料の統計解析ソフト**

- 自治医大さいたま医療センターの神田善伸先生が作成された

- **評価の高い無料統計ソフトRのGUI**

(Graphical User Interface : 情報の提示に画像や図形を多用し、基礎的な操作の大半を画面上の位置の指示により行うことができるような手法のこと)



EZRのインストール

インストールは下記サイトから：

<http://www.jichi.ac.jp/saitama-sct/SaitamaHP.files/statmed.html>

The screenshot shows the homepage of the Jichi Medical University Saitama Medical Center. The main content area features a central announcement for EZR version 1.41, dated December 1, 2019. It lists previous manual releases: the first manual in November 2014 and the second manual in April 2015. A new manual is announced for October 2016, titled 'EBM and Clinical Research'. The announcement includes a link to the English version and a note that the software is used in academic papers, with a reference to Bone Marrow Transplantation 2013: 48, 452-458. On the right side, there are Amazon.co.jp product listings for 'EBM and Clinical Research' and 'EZR and Statistical Analysis' manuals, both available for purchase.

自治医科大学附属さいたま医療センター
血液科のホームページへようこそ。

自治医科大学附属さいたま医療センター
Jichi Medical University Saitama Medical Center

ホーム | 施設紹介 | 移植診療 | 受診のご案内 | 研修・研究 | 統計ソフトEZR

フリー統計ソフトEZR

EZRの解析機能

EZRの使い方、変更履歴

ダウンロード (Windows標準版)

ダウンロード (MacOS X版)

ダウンロード (LINUX版)

よくあるご質問 (FAQ)

English version

***** Click here for English version *****

639280
(2012年6月11日のページ改訂後のアクセス数です。)

2019年12月1日 EZR version 1.41公開

2014年11月初心者向けマニュアル刊行

2015年4月EZRMニュアル第2版刊行
(↑クリックすると立ち読みすることができます)

2016年10月みんなのEBMと臨床研究刊行
(↑EBMの基礎、新倫理指針に連応した臨床研究の進め方、論文の書き方をまとめています)

1.23からサンプルの背景データのサマリー表(Table 1)を自動作成する新機能を搭載!!

1.40から傾向スコアマッチングなので必要なキャリパーマッチングに対応!!
動作確認OS Windows XP~10、Mac OS X Snow Leopard~Mojave、Ubuntu 11.10~15.10

EZRを使用した学術論文を発表される場合は
Bone Marrow Transplantation 2013; 48, 452-458
を参考文献として引用くださいますようお願いいたします。
(「EZRの使い方、変更履歴」を参照。2019年11月の時点で英文論文での引用回数が3000回を超えました)

EBMと臨床研究の解説書
amazon.co.jp
ゼロから始めて一冊でわかる...
¥ 3,960 Prime
今すぐ購入

EZRと統計解析の解説書
amazon.co.jp
初心者ですぐにできるフリ...
¥ 4,180 Prime
今すぐ購入

うまくいかない場合はFAQ

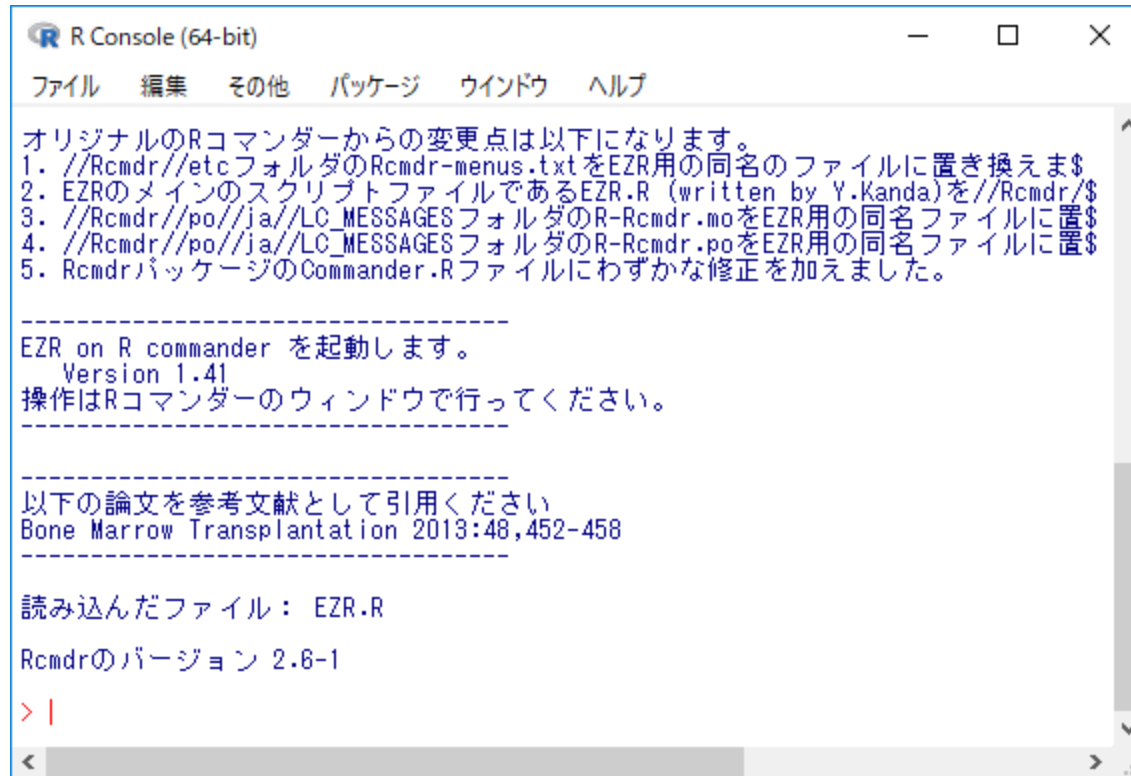
- どうしてもうまくいかない場合は、FAQ
<http://www.jichi.ac.jp/saitama-sct/SaitamaHP.files/statmedFAQ.html>
などご参照ください。
- EZRのダウンロード・インストールに関連して生じた一切の損害について、責任を負いかねますので、十分な注意の上、操作を行って下さい。
- 古いバージョンのRまたはEZRがPCにインストールされている場合、アンインストールしてから、改めてインストールする必要があります。

インストール後



デスクトップ上の
アイコンを
ダブルクリック

R Console



```
R Console (64-bit)
ファイル 編集 その他 パッケージ ウィンドウ ヘルプ

オリジナルのRコマンダーからの変更点は以下になります。
1. //Rcmdr//etcフォルダのRcmdr-menus.txtをEZR用の同名のファイルに置き換えま$
2. EZRのメインのスク립トファイルであるEZR.R (written by Y.Kanda)を//Rcmdr/$
3. //Rcmdr//po//ja//LC_MESSAGESフォルダのR-Rcmdr.moをEZR用の同名ファイルに置$
4. //Rcmdr//po//ja//LC_MESSAGESフォルダのR-Rcmdr.poをEZR用の同名ファイルに置$
5. RcmdrパッケージのCommander.Rファイルにわずかな修正を加えました。

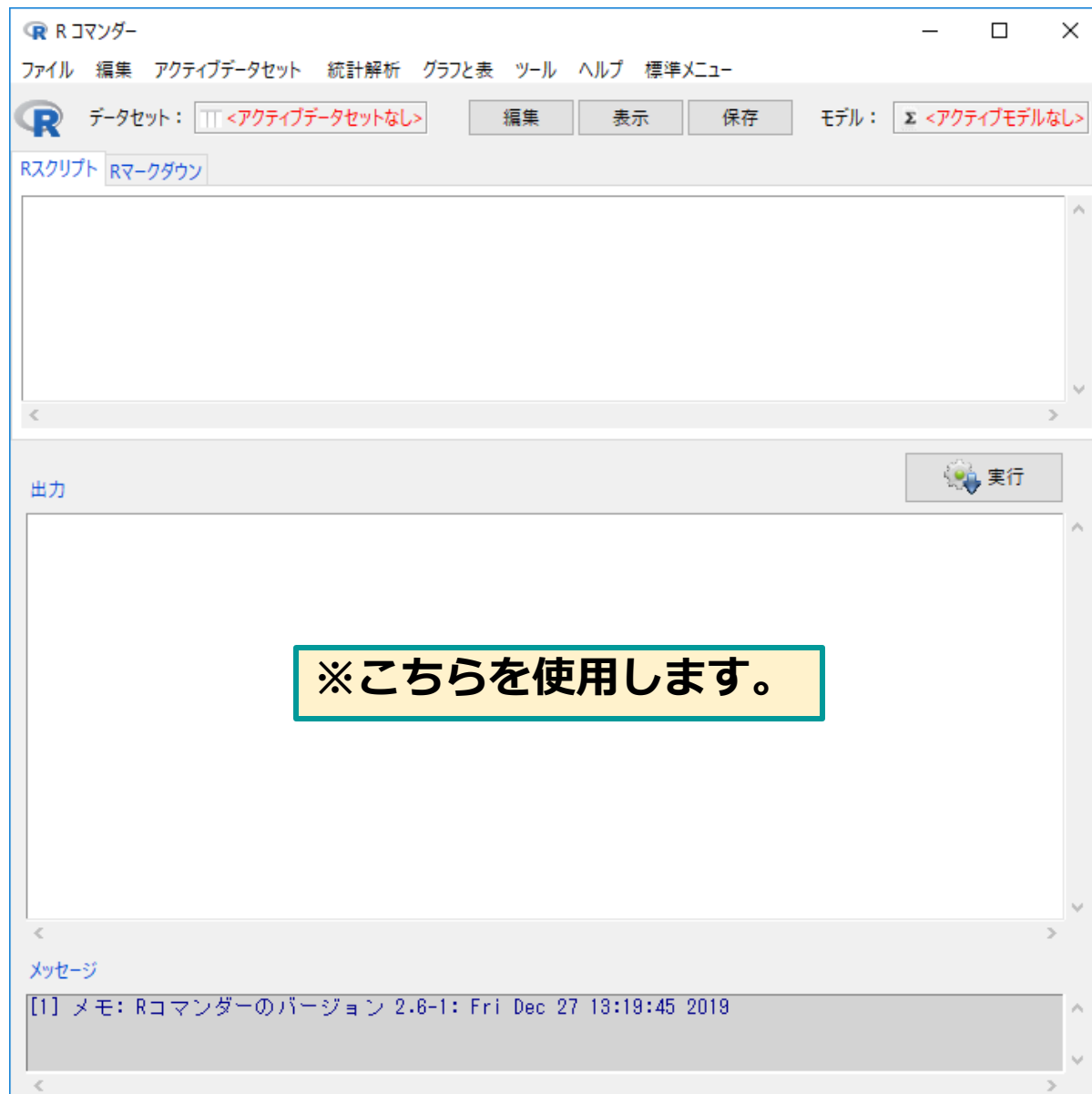
-----
EZR on R commander を起動します。
  Version 1.41
操作はRコマンダーのウィンドウで行ってください。
-----

以下の論文を参考文献として引用ください
Bone Marrow Transplantation 2013;48,452-458
-----

読み込んだファイル: EZR.R
Rcmdrのバージョン 2.8-1
> |
```

※こちらは使用しませんが、消さないでください。

R コマンダー



注意: Mac Userは少々大変です

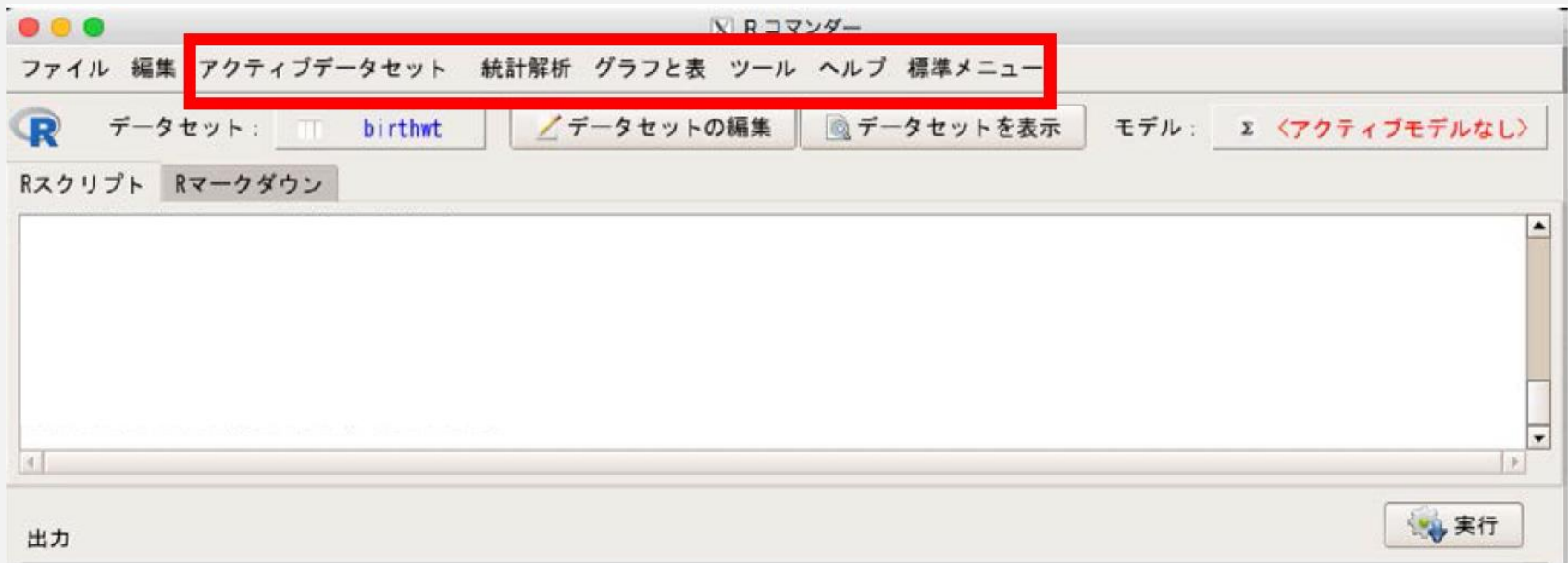
- Mac へのインストールは若干複雑ですが、下記サイトに記載されているとおりの順番で行ってください。

<http://www.jichi.ac.jp/saitama-sct/SaitamaHP.files/statmedOSX.html>

- どうしてもうまく行かない場合は、
別添のPDF資料「EZRのダウンロード_Mac.pdf」
もご参照いただきながら、インストールを実施してください。

注意: Mac Userは少々大変です

- Macでは、EZRがインストールできた場合、以下の画面が出てきます。赤枠の中が同じことを確認してください。



注意: Mac Userは少々大変です

- 当日のWebinarは、Windows versionを使った演習のため、Macとは若干見た目が異なります。
- 当日の演習でご不明な点がございましたら、ライブイベントのQ&A機能を用いてご質問ください。

2.2 事前学習: サンプルデータの作成と取り込み

早速統計処理、その前に

- 統計処理はデータあってのもの
- サンプルデータを作成してみましよう

データの型

連続型

年齢
身長
体重
検査値
スコア

離散型

(カテゴリー・名義など)

順位あり

- ・ ステージ、クラス

順位なし

- ・ 性別、合併症

データの型に応じて処理方法が異なる

⇒ 最初から区別することが大切

データファイルの作り方 1

大原則

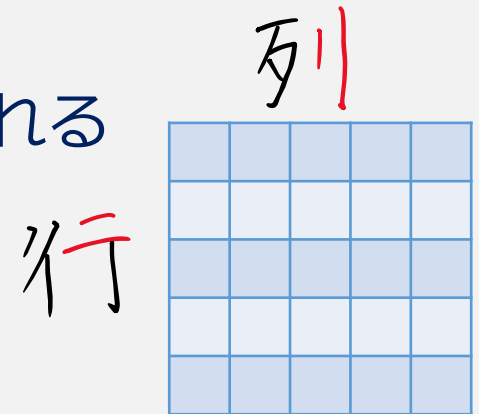
ひとつのセルにひとつの情報を

行：ひとりの患者さん（独立な測定単位）

列：変数（性別、年齢、検査値など）

□ ほとんどの統計ソフトの決まり事

- 統計処理は列方向の選択に対して行われる
- 1行目は変数名（半角英数字を推奨）



データファイルの作り方 2

□ 個人名ではなく**識別番号**で。

□ **連続型**

- すべて**半角数値**を使う（少数点も）。
- 「89.3」という数値の場合、
「89.3」「89.3」「>89」「89以上」などはダメ。
- 半角数値以外が入ると文字列として認識されるため、解析できない。

測定限界を超えた値は統計家と相談し
変数の入力方法を統一してください。

例えば

0.2未満→0.1（測定限界と0の中央値）

2000以上→2001 など

□ 欠測は空欄か「NA」などに**統一**しておく。

□ 日付は形式を**統一**しておく。

- YYYY-MM-DD、YYYY/MM/DD、YY-MM-DDなど。

適切な入力例

	A	B	C	D	E	F	G	H	I
1	ID	nday	type	treatmnt	prior	age	kps	censor	
2	10001	3	Adeno	Standard	0	43	30	0	
3	10002	7	Adeno	Test	0	58	40	0	
4	10003	8	Adeno	Standard	NA	61	20	1	
5	10004	8	Adeno	Test	0	66	50	0	
6	10005	12	Adeno	Standard	1	63	NA	1	
7	10006	18	Adeno	Test	1	69	NA	0	
8	10007	19	Adeno	Test	0	42	NA	0	
9	10008	24	Adeno	Test	0	60	40	1	

不適切な入力例

	A	B	C	D	E	F	G	H	I
1	ID	生存日数	セル	処置	前治療(あり1,なし0)	年齢	KPS	打ち切り	
2	10001	3	Adeno	Standard	0	43?	30	0	
3	10002	7	Adeno	Test	0	58	40	0	
4	10003	8	Adeno	標準	不明	61	20	1	
5	10004	8	Adeno	Test	0	66	50	0	
6	10005	12	Adeno	Standard	1	63		1	
7	10006	18	Adeno	Test	1	69		0	
8	10007	19	Adeno	Test	0	42		0	
9	10008	24	Adeno	Test	0	60	40	1	

データファイル作成の演習

以下の情報をもとに、従来薬に対する新薬の無作為化比較試験の**データファイル(csv形式)**を作成してみましょう。

- 成育一郎さんは、調査時、50歳、収縮期血圧140mmHg、降圧薬の服用があるかどうかは不明。割り付けを行ったところ、新薬であった。
- 次郎さんは、48歳、収縮期血圧140以上、降圧薬の服用があった。割り付けは従来薬であった。
- 三郎さんは、40歳、収縮期血圧120、降圧薬の服用はなかった。新薬に割り付けられた。

入力例

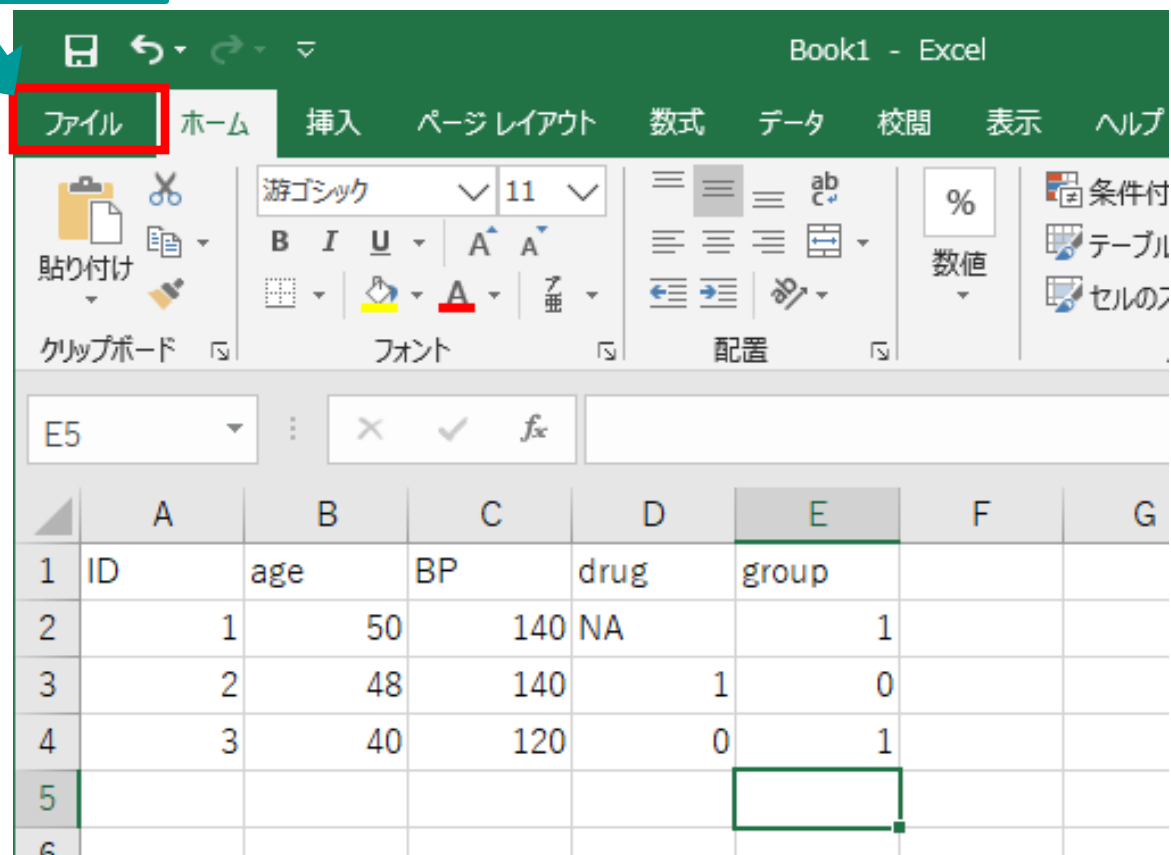
エクセルに以下のように入力してみましょう。

The screenshot shows the Microsoft Excel interface with the following data table:

	A	B	C	D	E	F	G
1	ID	age	BP	drug	group		
2	1	50	140	NA	1		
3	2	48	140	1	0		
4	3	40	120	0	1		
5							
6							

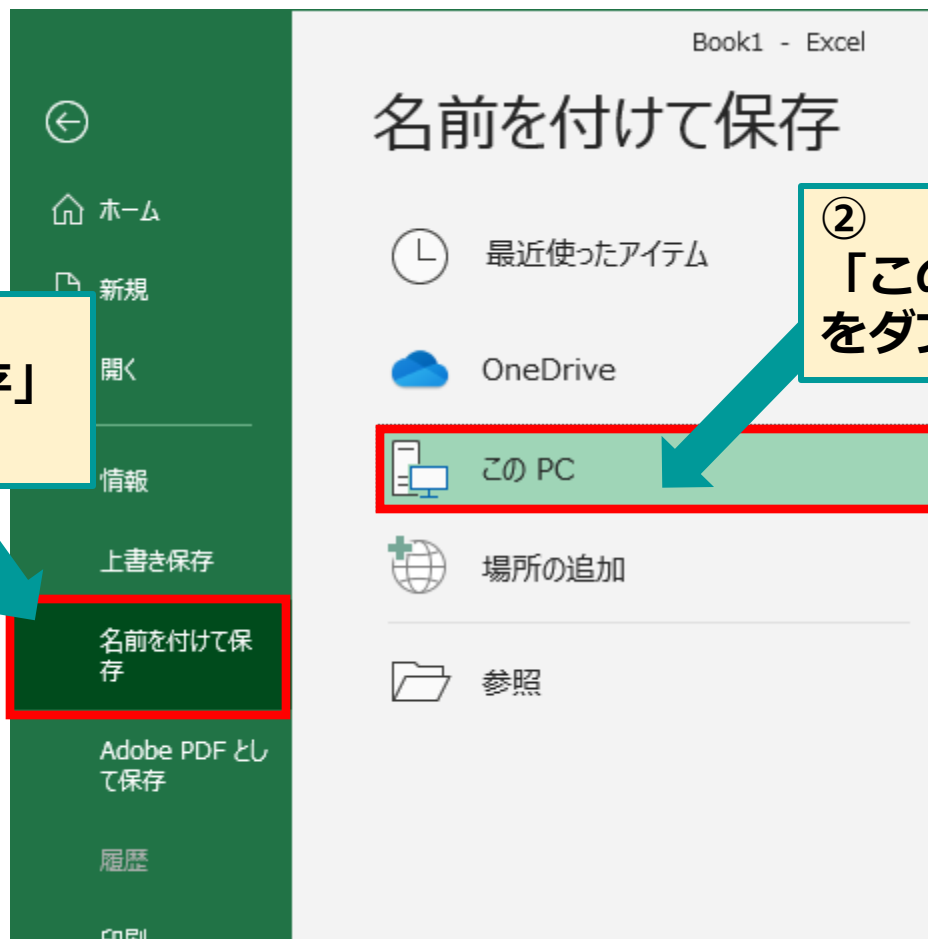
csv形式のファイルで保存①

ファイルをクリック



csv形式のファイルで保存②

①
「名前を付けて保存」
をクリック



②
「このPC」
をダブルクリック

csv形式のファイルで保存③

① 「デスクトップ」を選択



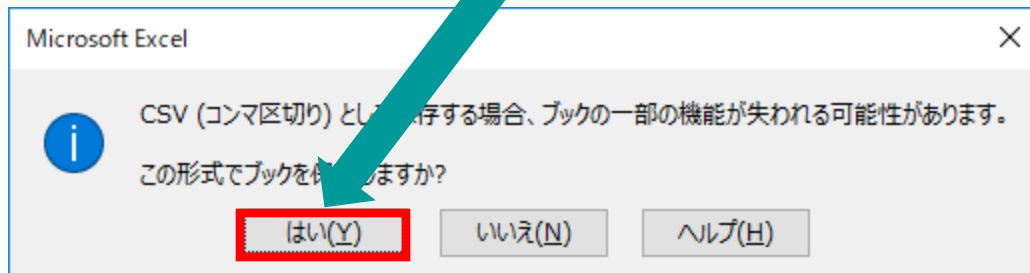
② ファイル名に「Sample」と入力

③ ファイルの種類を
「CSV (コンマ区切り)」に変更
※CSV UTF-8ではありません！

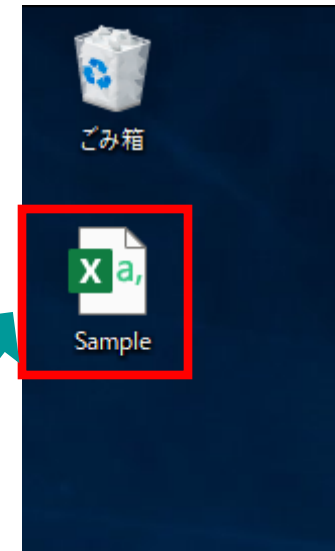
④ 「保存」をクリック

csv形式のファイルで保存④

① 「はい」 をクリック

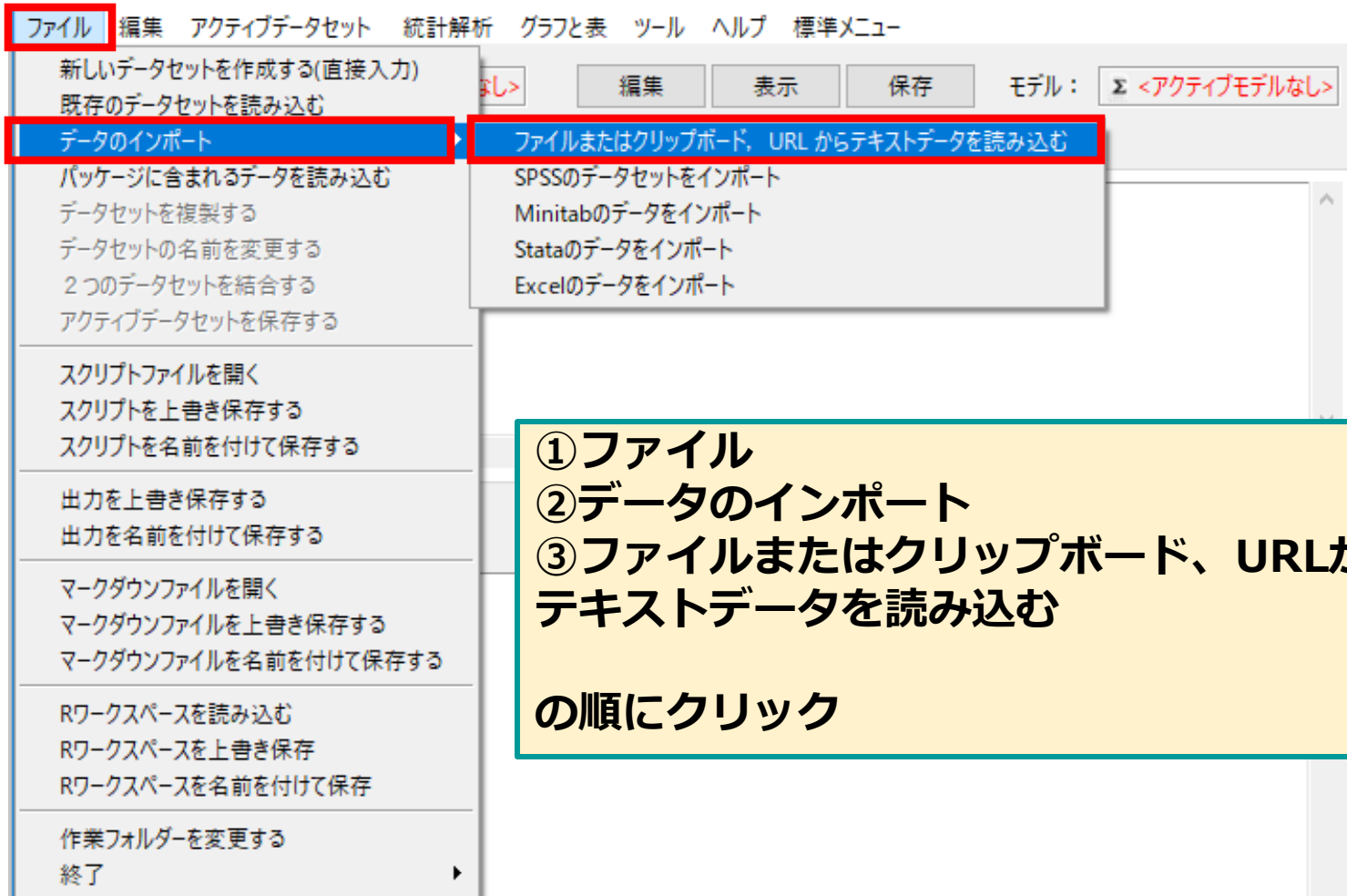


② デスクトップを確認



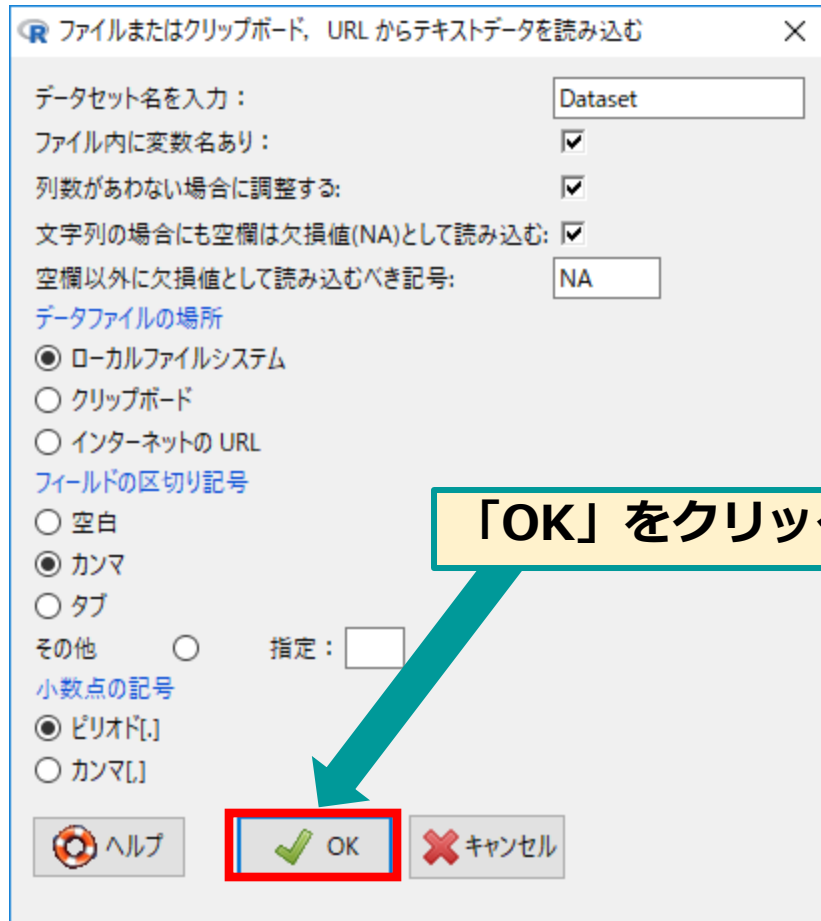


EZRでcsvファイルの取り込み①

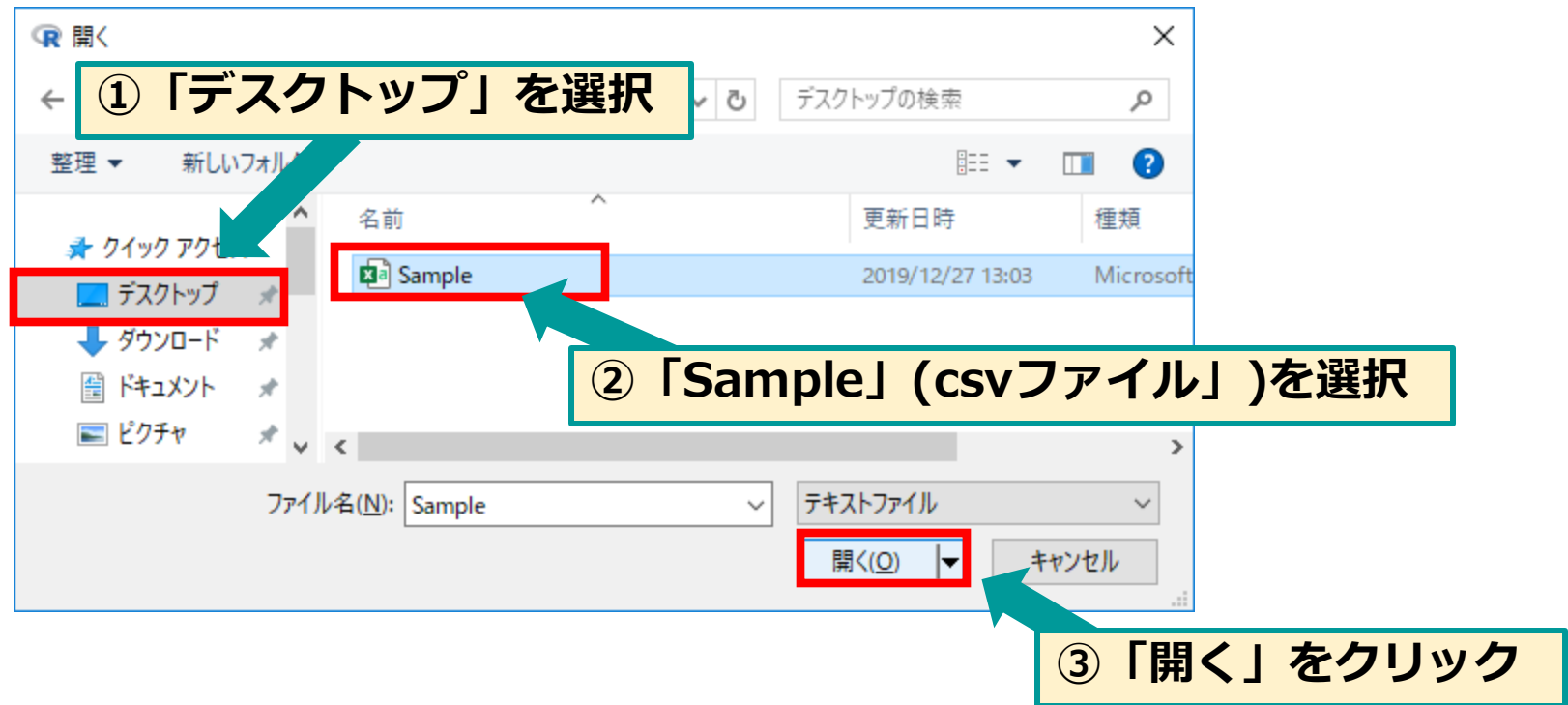


- ①ファイル
 - ②データのインポート
 - ③ファイルまたはクリップボード、URLからテキストデータを読み込む
- の順にクリック

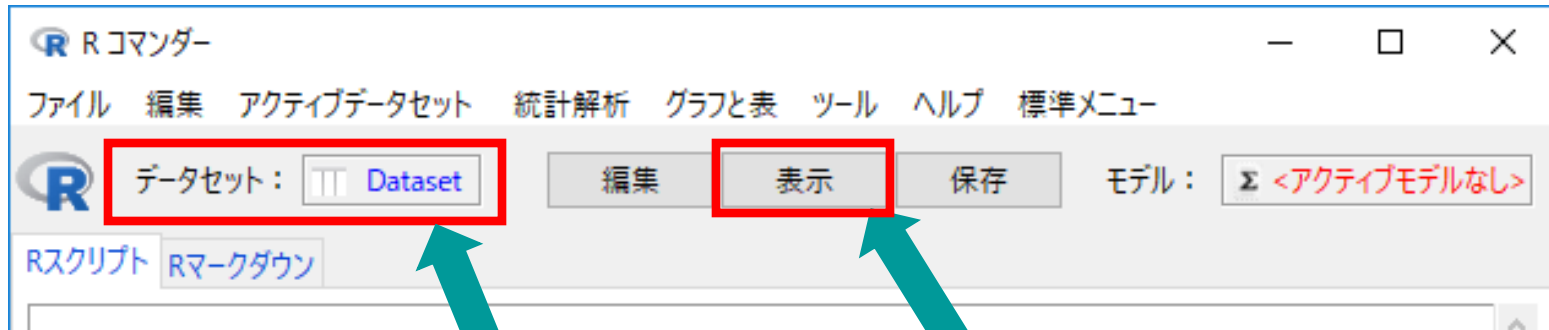
EZRでcsvファイルの取り込み②



EZRでcsvファイルの取り込み③



EZRでcsvファイルの取り込み④

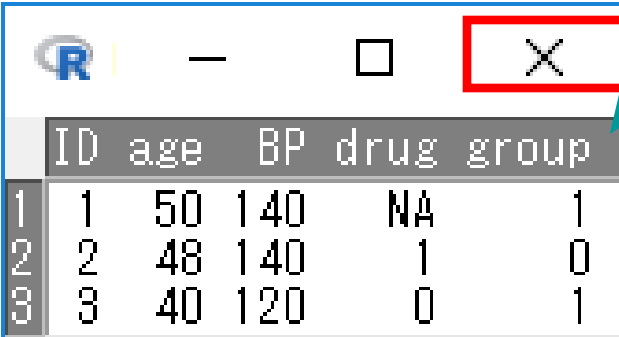


データセットが作成されると
青字になります

①「表示」をクリック

EZRでcsvファイルの取り込み⑤

うまくいくと表示されるデータ



	ID	age	BP	drug	group
1	1	50	140	NA	1
2	2	48	140	1	0
3	3	40	120	0	1

①確認できたら「×」（閉じる）をクリック

サンプルデータ NHANES

使用するサンプルデータ

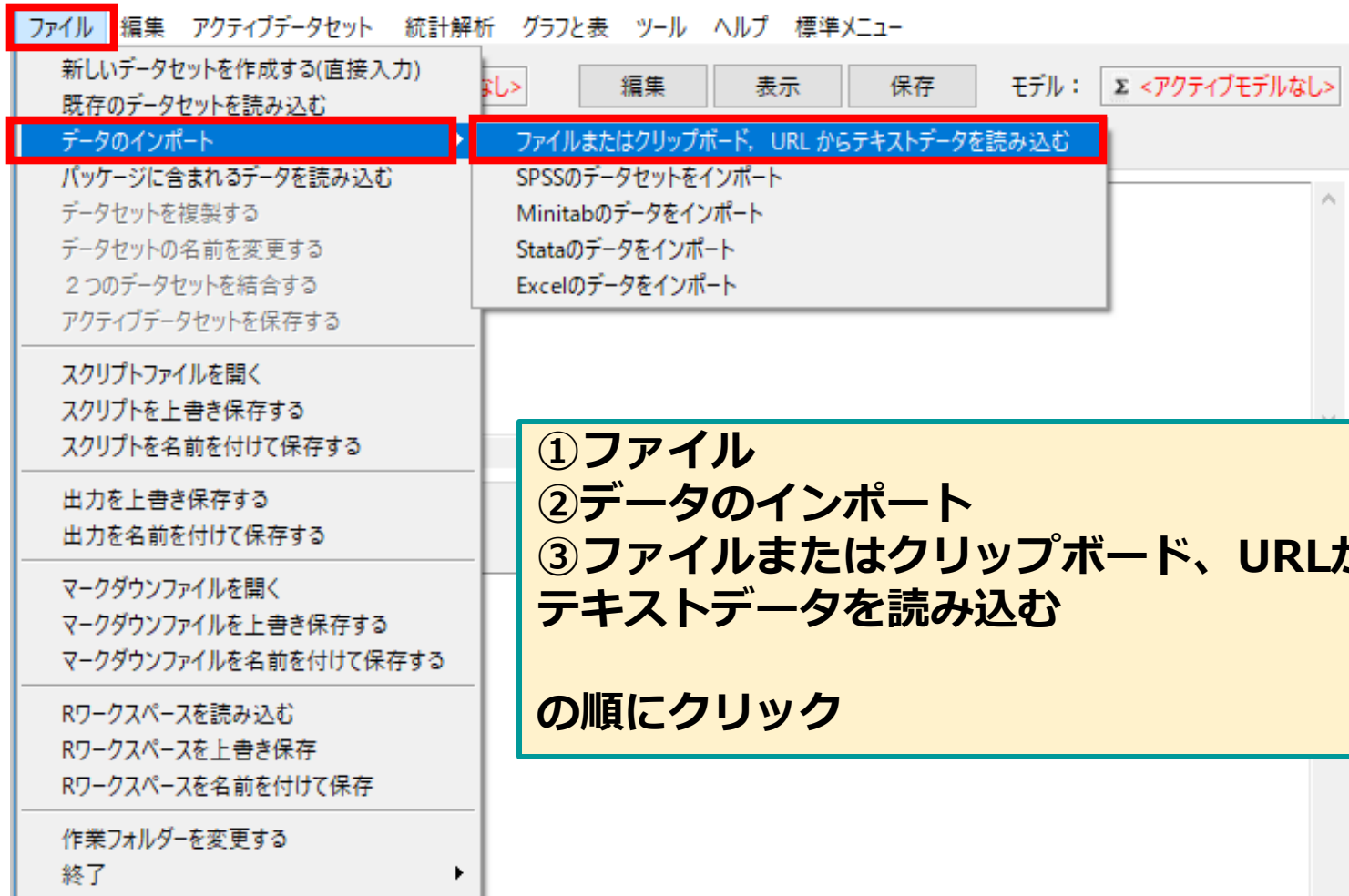
“NHANES” (演習用に編集したもの)を用いてcsvファイルのデータ取り込みを行います。

NHANES: 米国保健福祉省による全国健康栄養調査 “National Health and Nutrition Examination surveys”

事前にメールで送付した **NHANES.csv**をPCのデスクトップに保存してください。

EZRで、NHANES.csv を読み込んでみましょう。

NHANESデータの取り込み①



- ①ファイル
 - ②データのインポート
 - ③ファイルまたはクリップボード、URLからテキストデータを読み込む
- の順にクリック

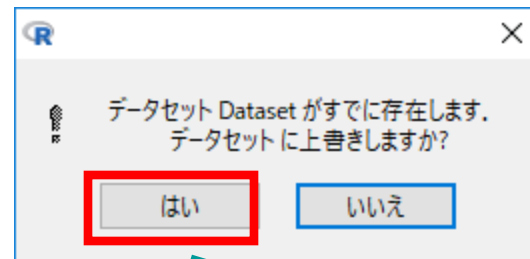
NHANESデータの取り込み②



EZR上で（アクティブ）データセットの名前を付けたいときは、ここを変更する。今回は変更しない。
（この場合、前に読み込んだ Sample.csv は、EZR上では上書きされて消去される。）

「OK」をクリック

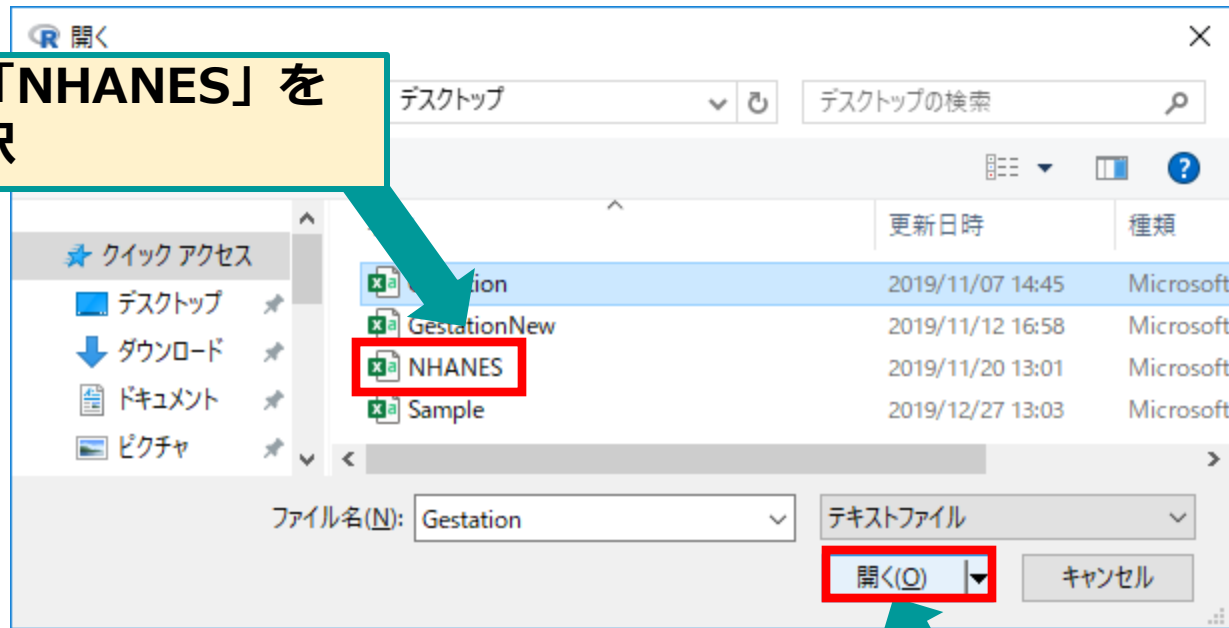
NHANESデータの取り込み③



「はい」をクリック

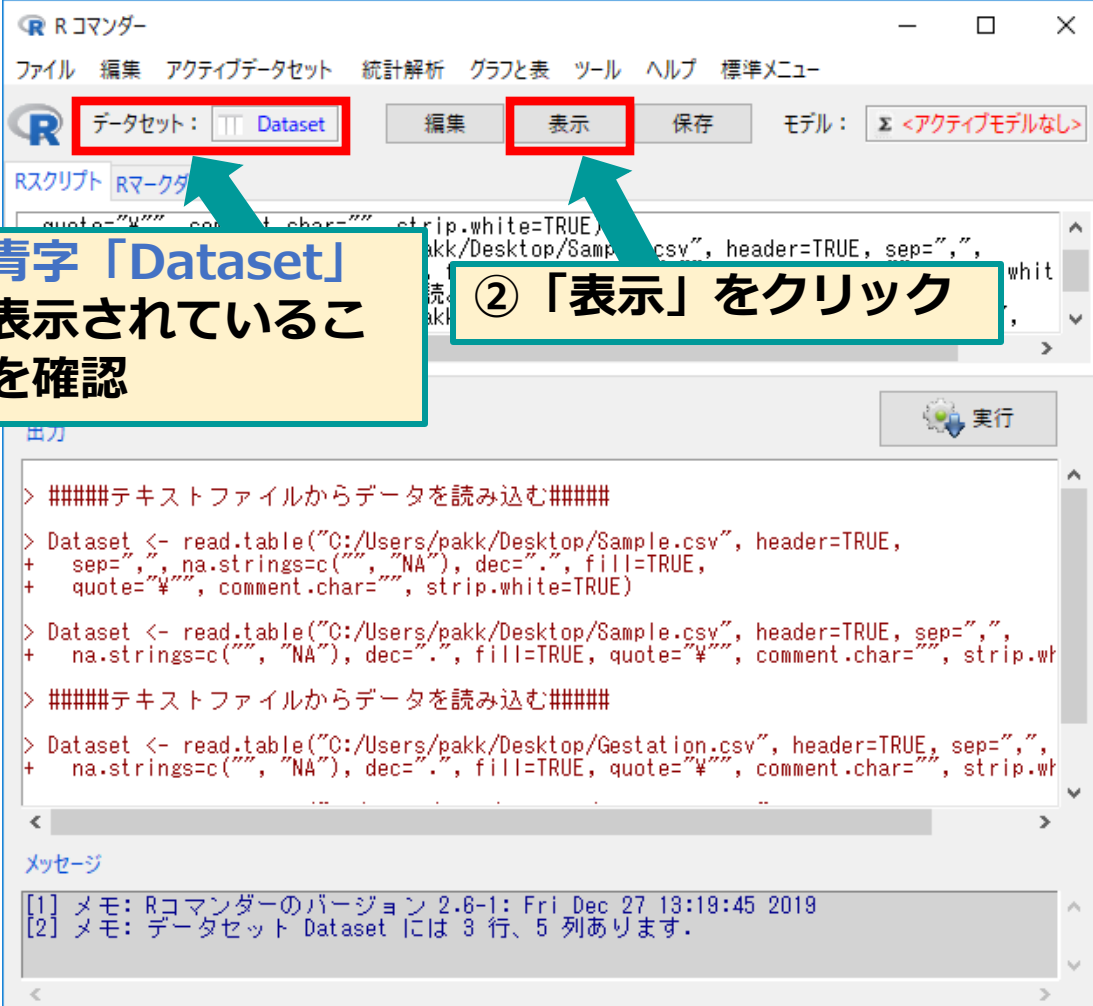
NHANESデータの取り込み④

① 「NHANES」を
選択



② 「開く」をクリック

NHANESデータの取り込み⑤



The screenshot shows the R Commander window with the following elements:

- Menu bar: ファイル 編集 アクティブデータセット 統計解析 グラフと表 ツール ヘルプ 標準メニュー
- Toolbar: データセット: Dataset (highlighted with a red box), 編集, 表示 (highlighted with a red box), 保存, モデル: <アクティブモデルなし>
- Code editor: Contains R code for reading CSV files into a Dataset object.
- Output window: Shows the execution of the R code.
- Message window: Shows system messages.

Two callout boxes provide instructions:

- ① 青字「Dataset」が表示されていることを確認 (Confirm that the word "Dataset" is displayed in blue)
- ② 「表示」をクリック (Click "表示")

```
> #####テキストファイルからデータを読み込む#####  
> Dataset <- read.table("C:/Users/pakk/Desktop/Sample.csv", header=TRUE,  
+ sep=";", na.strings=c("", "NA"), dec=".", fill=TRUE,  
+ quote="\"", comment.char="", strip.white=TRUE)  
> Dataset <- read.table("C:/Users/pakk/Desktop/Sample.csv", header=TRUE, sep=";",  
+ na.strings=c("", "NA"), dec=".", fill=TRUE, quote="\"", comment.char="", strip.wh  
> #####テキストファイルからデータを読み込む#####  
> Dataset <- read.table("C:/Users/pakk/Desktop/Gestation.csv", header=TRUE, sep=";",  
+ na.strings=c("", "NA"), dec=".", fill=TRUE, quote="\"", comment.char="", strip.wh
```

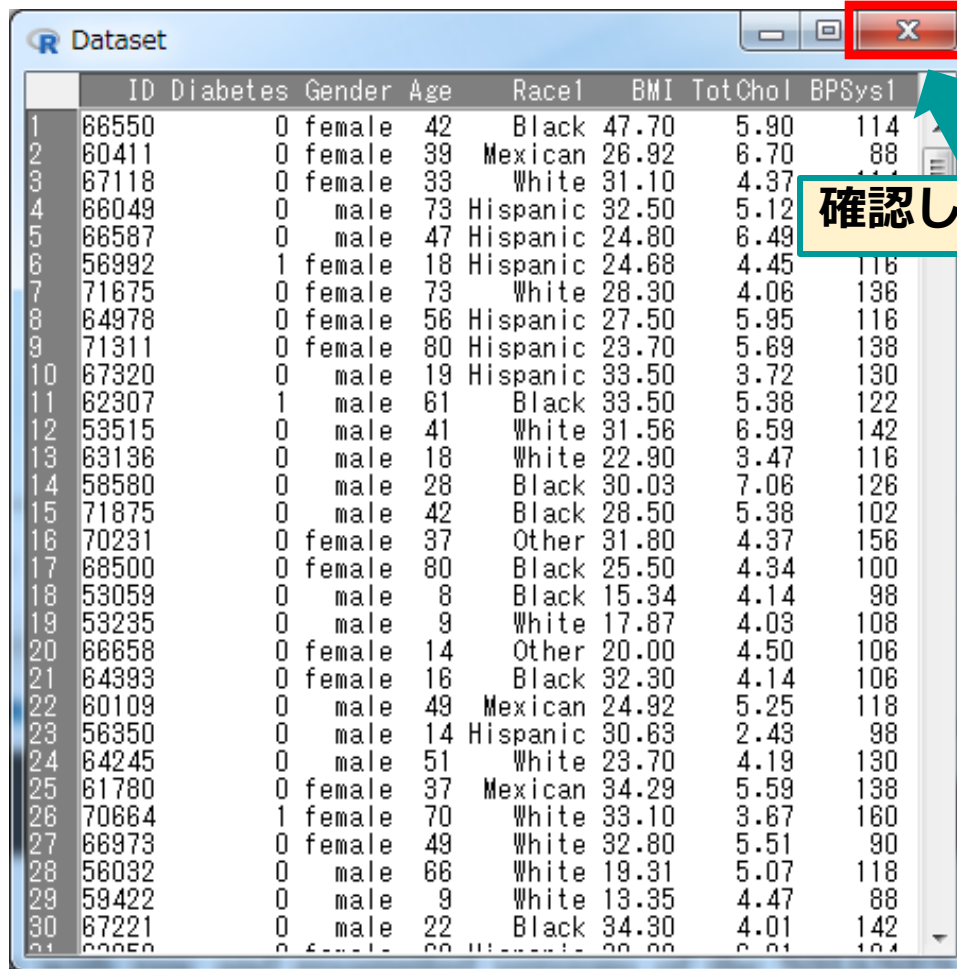
出力

```
> #####テキストファイルからデータを読み込む#####  
> Dataset <- read.table("C:/Users/pakk/Desktop/Sample.csv", header=TRUE,  
+ sep=";", na.strings=c("", "NA"), dec=".", fill=TRUE,  
+ quote="\"", comment.char="", strip.white=TRUE)  
> Dataset <- read.table("C:/Users/pakk/Desktop/Sample.csv", header=TRUE, sep=";",  
+ na.strings=c("", "NA"), dec=".", fill=TRUE, quote="\"", comment.char="", strip.wh  
> #####テキストファイルからデータを読み込む#####  
> Dataset <- read.table("C:/Users/pakk/Desktop/Gestation.csv", header=TRUE, sep=";",  
+ na.strings=c("", "NA"), dec=".", fill=TRUE, quote="\"", comment.char="", strip.wh
```

メッセージ

```
[1] メモ: Rコマンドのバージョン 2.6-1: Fri Dec 27 13:19:45 2019  
[2] メモ: データセット Dataset には 3 行、5 列あります。
```

NHANESデータの取り込み⑥



	ID	Diabetes	Gender	Age	Race1	BMI	TotChol	BPSys1
1	66550	0	female	42	Black	47.70	5.90	114
2	60411	0	female	39	Mexican	26.92	6.70	88
3	67118	0	female	33	White	31.10	4.37	114
4	66049	0	male	73	Hispanic	32.50	5.12	114
5	66587	0	male	47	Hispanic	24.80	6.49	114
6	56992	1	female	18	Hispanic	24.68	4.45	116
7	71875	0	female	73	White	28.30	4.06	136
8	64978	0	female	56	Hispanic	27.50	5.95	116
9	71311	0	female	80	Hispanic	23.70	5.69	138
10	67320	0	male	19	Hispanic	33.50	3.72	130
11	62307	1	male	61	Black	33.50	5.38	122
12	53515	0	male	41	White	31.56	6.59	142
13	63136	0	male	18	White	22.90	3.47	116
14	58580	0	male	28	Black	30.03	7.06	126
15	71875	0	male	42	Black	28.50	5.38	102
16	70231	0	female	37	Other	31.80	4.37	156
17	68500	0	female	80	Black	25.50	4.34	100
18	53059	0	male	8	Black	15.34	4.14	98
19	53235	0	male	9	White	17.87	4.03	108
20	66658	0	female	14	Other	20.00	4.50	106
21	64393	0	female	16	Black	32.30	4.14	106
22	60109	0	male	49	Mexican	24.92	5.25	118
23	56350	0	male	14	Hispanic	30.63	2.43	98
24	64245	0	male	51	White	23.70	4.19	130
25	61780	0	female	37	Mexican	34.29	5.59	138
26	70664	1	female	70	White	33.10	3.67	160
27	66973	0	female	49	White	32.80	5.51	90
28	56032	0	male	66	White	19.31	5.07	118
29	59422	0	male	9	White	13.35	4.47	88
30	67221	0	male	22	Black	34.30	4.01	142
31	66550	0	female	42	Black	47.70	5.90	114

確認したら閉じる

NHANESデータの変数一覧

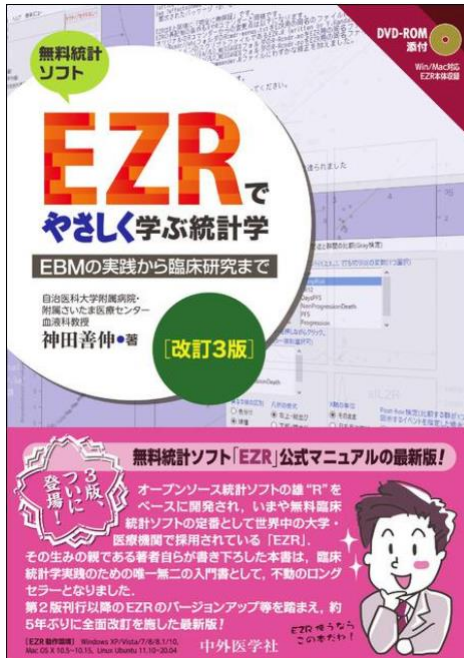
NHANESデータの変数一覧

変数名	項目名	データの型	カテゴリー区分
ID	参加者ID	離散	
Diabetes	糖尿病の有無	連続 離散	1: Yes, 0: No
Gender	性別	連続 離散	male or female
Age	年齢	連続 離散	
Race1	人種	連続 離散	Mexican, Hispanic, White, Black, or Other.
BMI	BMI (kg/m ²)	連続 離散	
TotChol	HDLコレステロール値 (mmol/L)	連続 離散	
BPSys1	収縮期血圧 (mmHg)	連続 離散	

受講者の皆様にお問い合わせする事前準備は以上です。

当日、皆様とご一緒できることを
楽しみにしております！！

参考文献



**EZRでやさしく学ぶ統計学 改訂3版
～EBMの実践から臨床研究まで～**
神田善伸 著

<http://www.chugaiigaku.jp/item/detail.php?id=3342>



**サラッとできる！フリー統計ソフトEZ R
(Easy R) でカンタン統計解析**
神田善伸 著

<https://www.ohmsha.co.jp/book/9784274226328/>