

事前準備作業

Rの環境設定

はじめに

Rの環境設定と設定確認に関する説明資料になります。
研修前に一読いただき設定して頂きますようお願い致します。

目次

- | | |
|------------------|----------------------|
| 1.Rのインストール手順 | 7.外部データファイルの読み込み |
| 2.パッケージのインストール準備 | 8.データフレームの概要 |
| 3.R画面の起動 | 9.R画面の終了 |
| 4.パッケージのインストール | 10.R利用の基礎 |
| 5.作業フォルダの設定 | 11.注意事項 |
| 6.フォントの設定 | 12.（お願い）研修資料の取扱いについて |

1.Rのインストール手順（1/6）

Windows版のRのインストールするため、以下のURLへアクセスして下さい。

<https://cran.r-project.org/bin/windows/base/old/4.1.3/>

Index of /bin/windows/base/old/4.1.3

Name	Last modified	Size	Description
 Parent Directory	-		
 NEWS.R-4.1.3.html	2022-03-10 09:16	109K	
 R-4.1.3-win.exe	2022-03-10 10:32	86M	
 README.R-4.1.3	2022-03-10 09:16	8.5K	
 SVN-REVISION.R-4.1.3	2022-03-10 09:16	46	
 md5sum.txt.R-4.1.3	2022-03-10 10:32	50	
 release.html	2022-03-10 09:16	90	
 rw-FAQ.html	2022-03-10 09:16	88K	

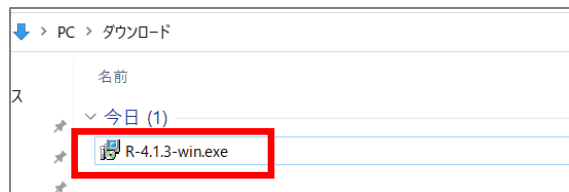
Apache Server at cran.r-project.org Port 443

1

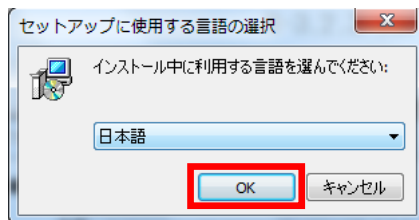
「R 4.1.3-win.exe」をクリックしダウンロードする。

最新バージョンではありませんが、
本研修はRバージョン「4.1.3」を
利用して下さい

1.Rのインストール手順（2/6）

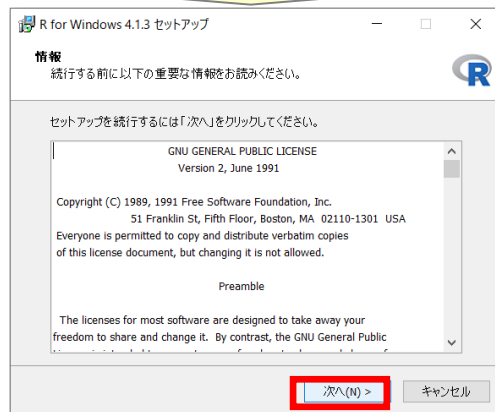


2 保存先のフォルダを開き「R-4.1.3-win.exe」
ファイルをダブルクリックする



3 【セットアップに使用する言語の選択】
「OK」ボタンを押下する。
※利用する言語は「日本語」

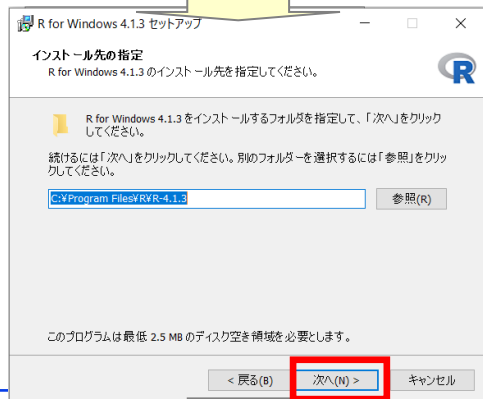
1.Rのインストール手順 (3/6)



4

【情報】

「次へ」ボタンを押下する。

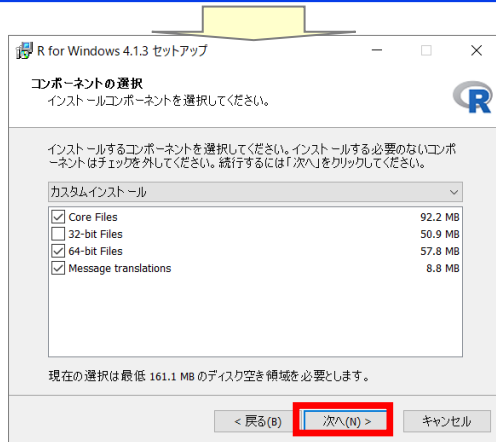


5

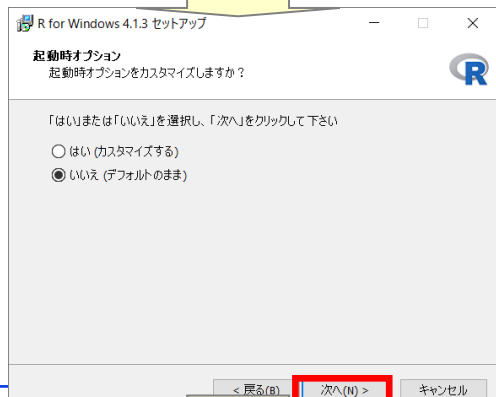
【インストール先の指定】

「次へ」ボタンを押下する。
※インストール先を特に変更
する必要がない場合は、デ
フォルトのままでもよい。

1.Rのインストール手順（4/6）

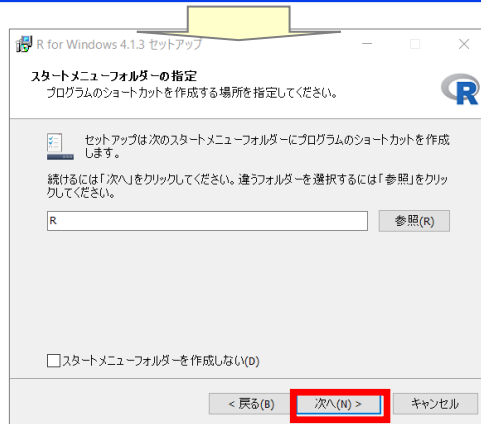


6 **【コンポーネントの選択】**
日本語のメッセージが必要な場合はデフォルトの設定の他「Message translations」にチェックを入れ、「次へ」ボタンを押下する。



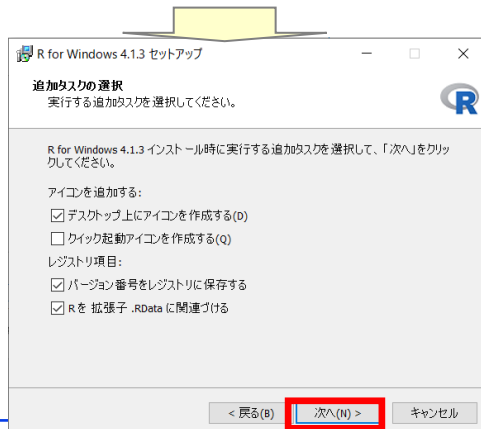
7 **【起動時オプション】**
「いいえ」を選択のまま、「次へ」ボタンを押下する。

1.Rのインストール手順（5/6）



8

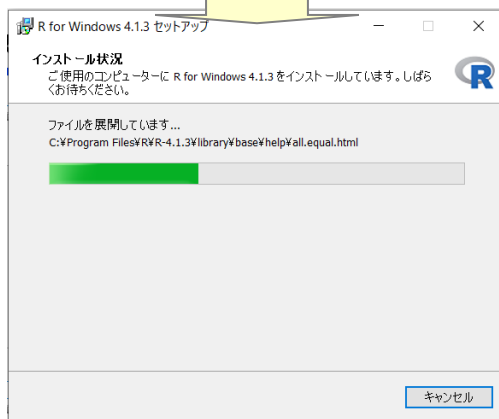
【スタートメニューフォルダの設定】
デフォルトのまま「次へ」ボタンを押下する。



9

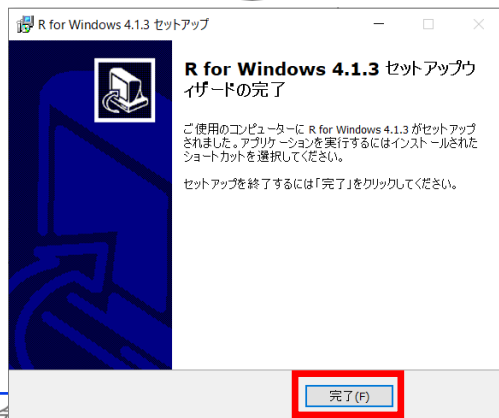
【追加タスクの選択】
「デスクトップ上にアイコンを作成する」にチェックし、その他のチェックはデフォルトのまま「次へ」ボタンを押下する。

1.Rのインストール手順（6/6）



10

【インストール状況】
インストールが開始される。

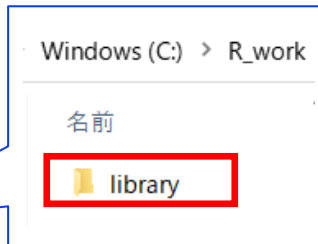
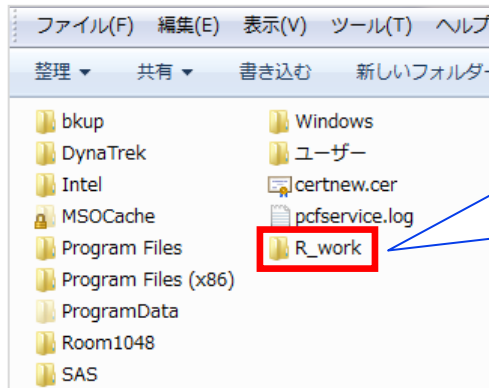


11

【R for windows 4.1.3 セットアップウィザードの完了】
インストールが終了すると完了画面が表示されるので、「完了」ボタンを押下する。

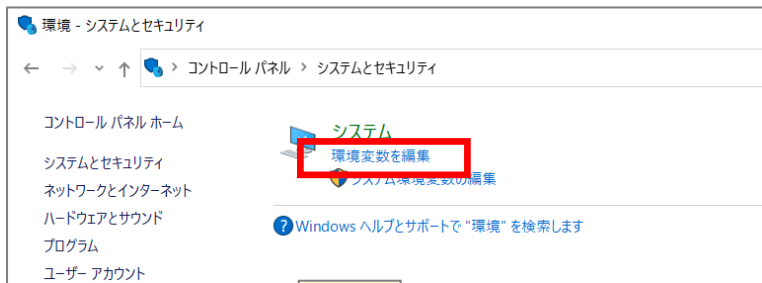
2.パッケージのインストール準備 (1/3)

通常パッケージで追加する場合はマイドキュメント配下に追加されます。
ファイル削除バッチで削除されない為にライブラリ追加先を変更します。



1

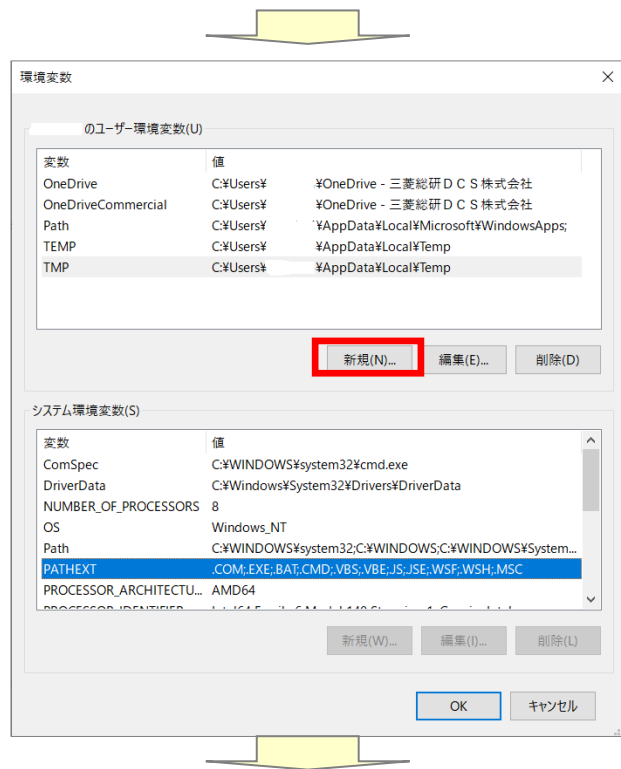
Cドライブの直下に、研修用作業フォルダ「R_work」を作成し、その直下に「library」フォルダを作成する



2

コントロールパネルの「システムとセキュリティ」にある「環境変数を編集」をクリックする。

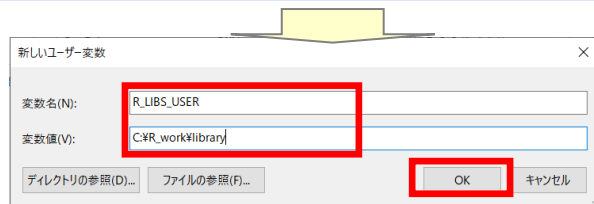
2.パッケージのインストール準備 (2/3)



3

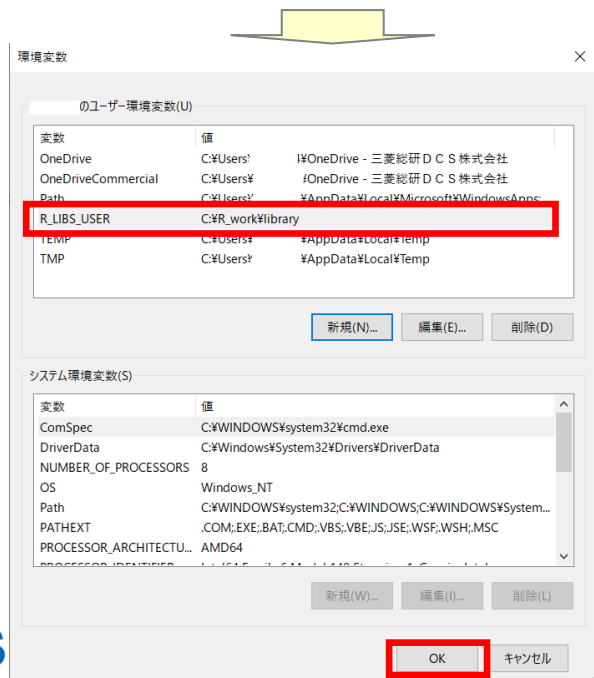
【環境変数】
「新規」ボタンをクリックする。

2.パッケージのインストール準備 (3/3)



4

【新しいユーザー変数】
下記とおりに入力して、「OK」ボタンをクリックする
変数名 R_LIBS_USER
変数値 C:\R_work\library

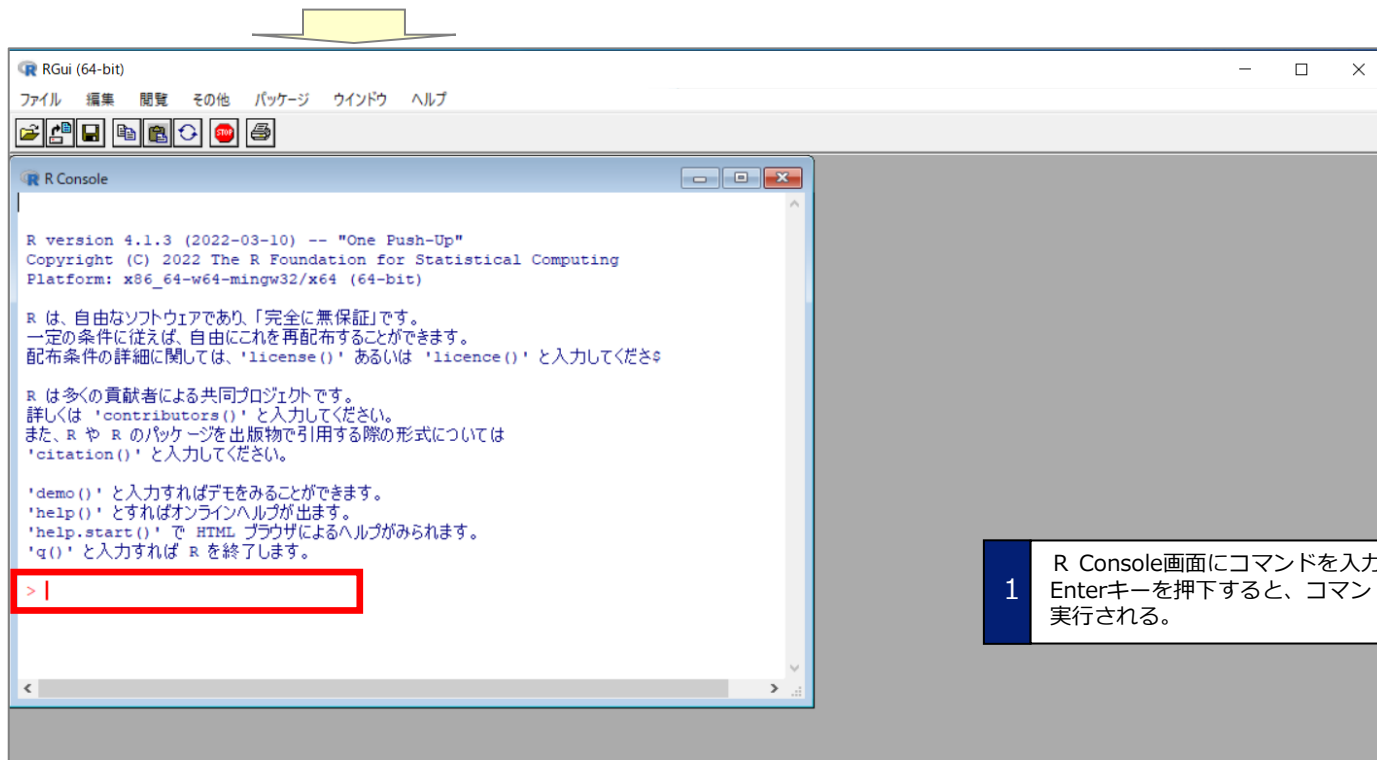


5

【環境変数】
環境変数ウィンドウの“（自身のユーザー名）のユーザー環境変数”に新しく環境変数が追加されていることを確認し、「OK」ボタンをクリックする。

3. R画面の起動

インストールしたRを起動します。



1

R Console画面にコマンドを入力して、Enterキーを押下すると、コマンドが実行される。

4. パッケージのインストール (1/4)

Rはインストールした初期状態でも標準的な統計手法を利用することができますが、さらに様々な統計手法を使えるようにするために、「パッケージ」と呼ばれるものが用意されています。

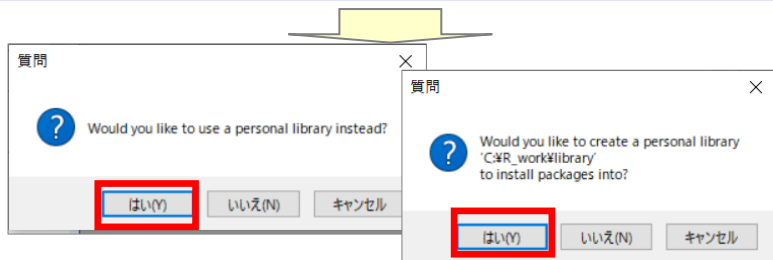
「パッケージ」のインストール手順について説明します。

1 「R Console」をアクティブにした状態で、メニュー「パッケージ」→「パッケージのインストール」を選択する。

2 「Secure CRAN mirror」ダイアログで「Japan (Tokyo)」を選択して「OK」をクリックする。

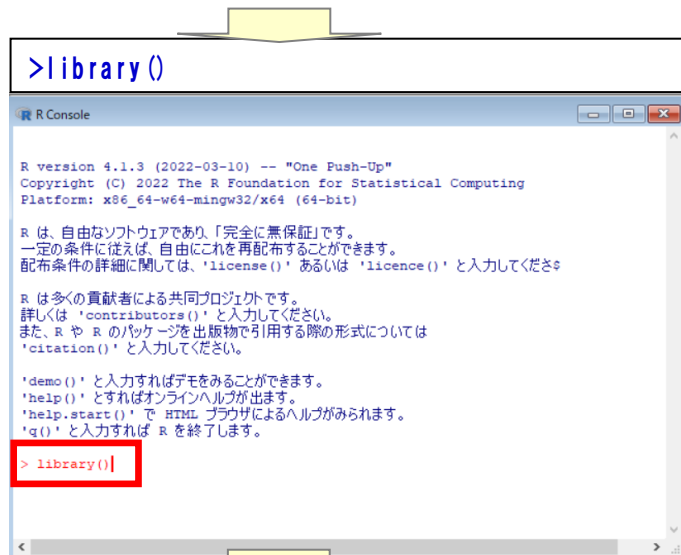
3 「Packages」ダイアログボックスのリストから**DAAG**を選択して「OK」をクリックする。

4. パッケージのインストール (2/4)



3

質問が表示された場合は「はい」をクリックする。



4

パッケージのインストール完了後、R Console画面にて「library()」コマンドを入力し実行し、利用可能なRパッケージの一覧が表示する。
DAAGパッケージが正常にインストールされていることを確認する。

パッケージ名	使用する関数	分析手法
DAAG	vif関数	重回帰分析（VIF値の算出）

4. パッケージのインストール (3/4)

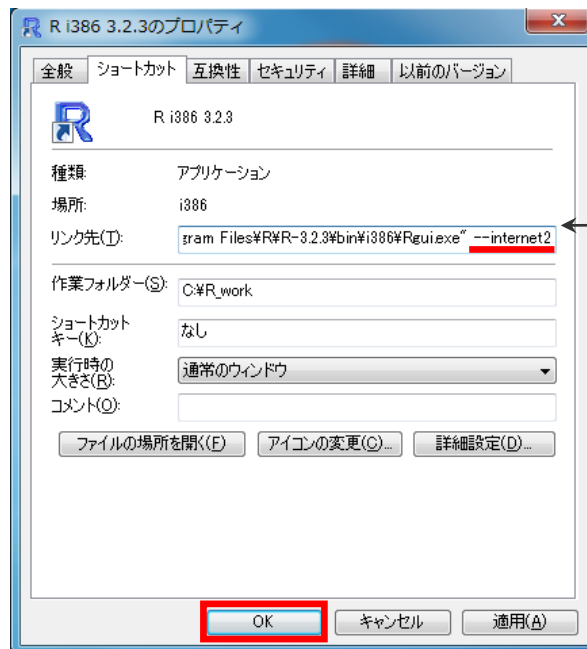
利用可能なRパッケージの一覧に、インストールしたパッケージが表示されていればOK。
表示されていない場合は、再度インストールを実施する。

層別ヒストグラムを作成する「lattice」パッケージがあることも併せて確認する。ない場合は、インストールを実施する。

4. パッケージのインストール (4/4)

パッケージのインストールができない場合、以下の設定を行ってください。

WindowsマシンでRを使用しており、かつプロキシサーバでインターネットにアクセスしている場合はRのプロパティの「リンク先」の後ろに「--internet2」を追記して保存後、再度パッケージのインストールを実施して下さい。

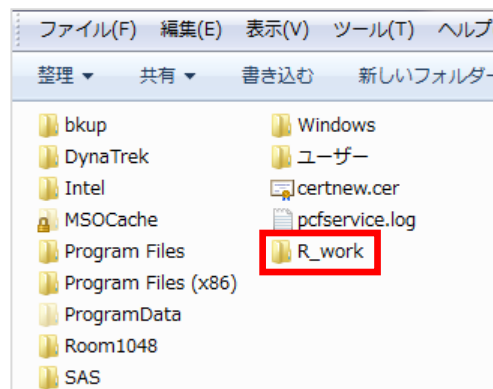


リンク先の後ろに
「--internet2」を追記
(--は半角。--の前に半角のスペースが必要)

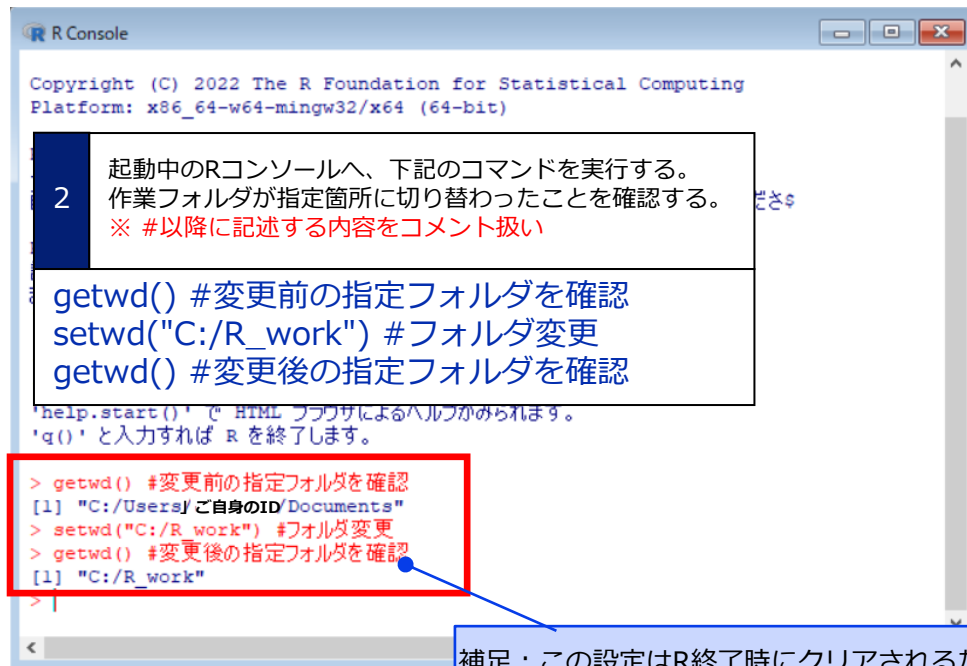
5.作業フォルダの設定 (1/2)

研修で使用する作業フォルダの設定を行います。

※研修で使用するサンプルデータは、「作業フォルダ」内へコピーして下さい。

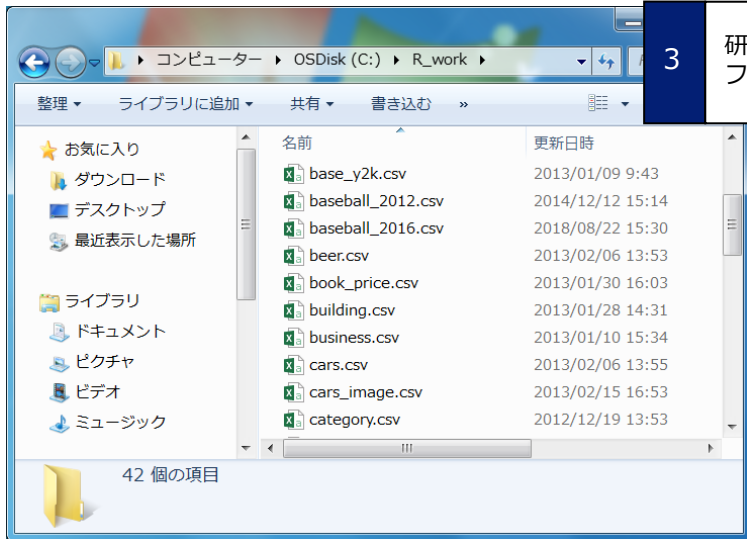


1 Cドライブの直下に、研修用作業フォルダ「R_work」を作成する



補足：この設定はR終了時にクリアされるため、R起動時に毎回実行が必要。

5.作業フォルダの設定 (2/2)



A screenshot of a Windows Explorer window showing the contents of the 'R_work' folder on the C: drive. The address bar shows the path 'コンピュータ > OSDisk (C:) > R_work'. The left sidebar shows the 'ライブラリ' (Libraries) section. The main pane displays a list of files with columns for '名前' (Name) and '更新日時' (Last Modified). The files listed are:

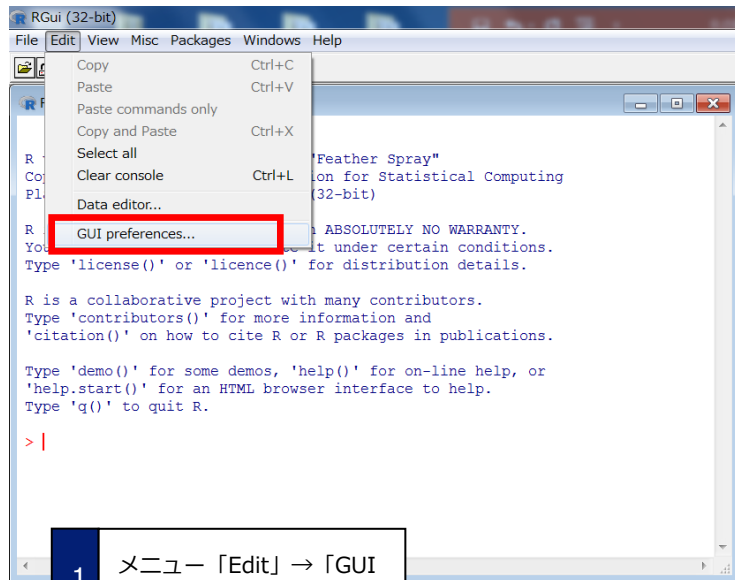
名前	更新日時
base_y2k.csv	2013/01/09 9:43
baseball_2012.csv	2014/12/12 15:14
baseball_2016.csv	2018/08/22 15:30
beer.csv	2013/02/06 13:53
book_price.csv	2013/01/30 16:03
building.csv	2013/01/28 14:31
business.csv	2013/01/10 15:34
cars.csv	2013/02/06 13:55
cars_image.csv	2013/02/15 16:53
category.csv	2012/12/19 13:53

The status bar at the bottom indicates '42 個の項目' (42 items). A yellow arrow points to the folder icon in the status bar.

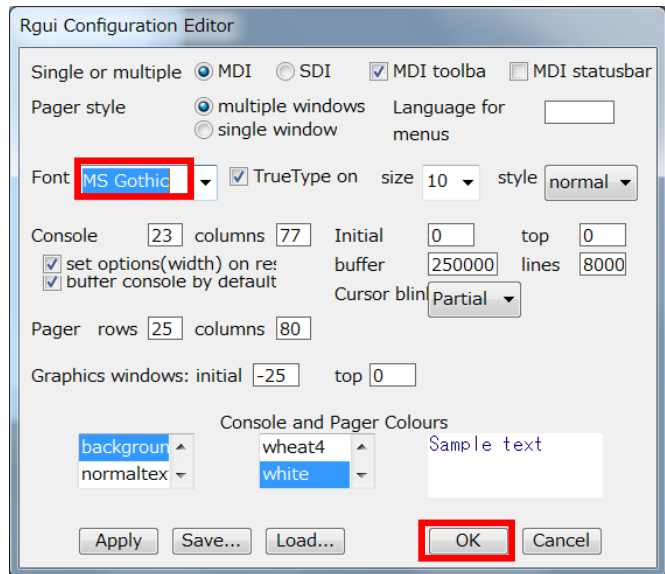
3 研修時に使用するサンプルデータを、作成したフォルダ「R_work」へコピーする。

6.フォントの設定

日本語を入力する際にカーソルがずれないように、フォントを設定します。



1 メニュー「Edit」→「GUI preferences」を選択する。



2 Rgui設定エディター画面で Fontを「MS Gothic」へ変更し OKボタンを押下する。

補足：この設定はR終了時にクリアされるため、R起動時に毎回設定が必要。

7.外部データファイルの読み込み

Rで野球データ (baseball_2016.csv) を変数「base」へ読み込み、「base」の中身を確認してみましょう。

```
>base=read.csv("baseball_2016.csv")  
>head(base)
```

【補足】
「=」の他に「<-」の記号も”右辺を左辺へ代入する”記号として使用できます。
※講義資料では、「=」の記号で記載してます。

出力結果

	REAGUE	TEAM	Team_Order	No	PLAYER POSITION		BIRTHDAY	AGE	ZAISEKI	NEN	HEIGHT	WEIGHT	BLOOD	TYPE	TOUDA
1	セ・リーグ	広島	1	0	中東 直己	外野手	1981/10/5	35	10	168	73		B型	右左	
2	セ・リーグ	広島	1	0	上本 崇司	内野手	1990/8/22	26	4	170	71		A型	右両	
3	セ・リーグ	広島	1	2	田中 広輔	内野手	1989/7/3	27	3	171	81		A型	右左	
4	セ・リーグ	広島	1	4	小窪 哲也	内野手	1985/4/12	31	9	175	80		O型	右右	
5	セ・リーグ	広島	1	5	ルナ	内野手	1980/2/1	36	4	187	99		B型	右右	
6	セ・リーグ	広島	1	6	梵 英心	内野手	1980/10/11	36	11	173	76		B型	右右	
	HOMETOWN	NEN_2016	NEN_2017	NEN_SAGAKU	UP_RATIO	KAIJYO	DARITSU	SHIAI	DASEKI	DASUU	TOKUTEN	ANDA	TWEBASE	THREEBASE	HR
1	広島	1300	NA	NA	NA	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
2	広島	880	700	-180	0.7954545	0	0.000	7	0	0	0	0	0	0	0
3	神奈川	4100	7800	3700	1.9024390	0	0.265	143	679	581	102	154	17	3	13
4	奈良	4000	4000	0	1.0000000	0	0.217	69	107	92	8	20	9	1	2
5	ドミニカ共和国	12000	NA	NA	NA	1	0.272	67	268	243	35	66	8	1	5
6	広島	9000	4000	-5000	0.4444444	0	0.000	7	12	9	1	0	0	0	0
	HIT	DATEN	TOURUI	TOURUI_SOSHI	GIDA	GISEIFURAI	FOUR_BALL	KEIEN	DEAD_BALL	SANSHIN	DOUBLE_PLAY	SLG	OPS	zougaku_flg	
1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	0	
2	0	0	0	0	0	0	0	0	0	0	0	0.000	0.000	0	
3	216	39	28	19	3	1	77	1	17	119	1	0.372	0.367	1	
4	37	10	0	0	1	0	11	0	3	23	2	0.402	0.321	0	
5	91	34	6	0	0	2	22	0	1	47	5	0.374	0.332	0	
6	0	0	0	0	0	0	3	0	0	0	2	0.000	0.250	0	

head関数はデータの最初の6行を表示させるコマンドです。

8.データフレームの概要

野球データ (baseball_2016.csv) を変数「base」へ読み込むと、データフレームが作成されます。データフレームでは、量的データと質的データの異なるデータタイプが混在していても扱うことができます。どのような型で読み込まれているか確認するためには、str関数を使用します。

```
>str(base)
```

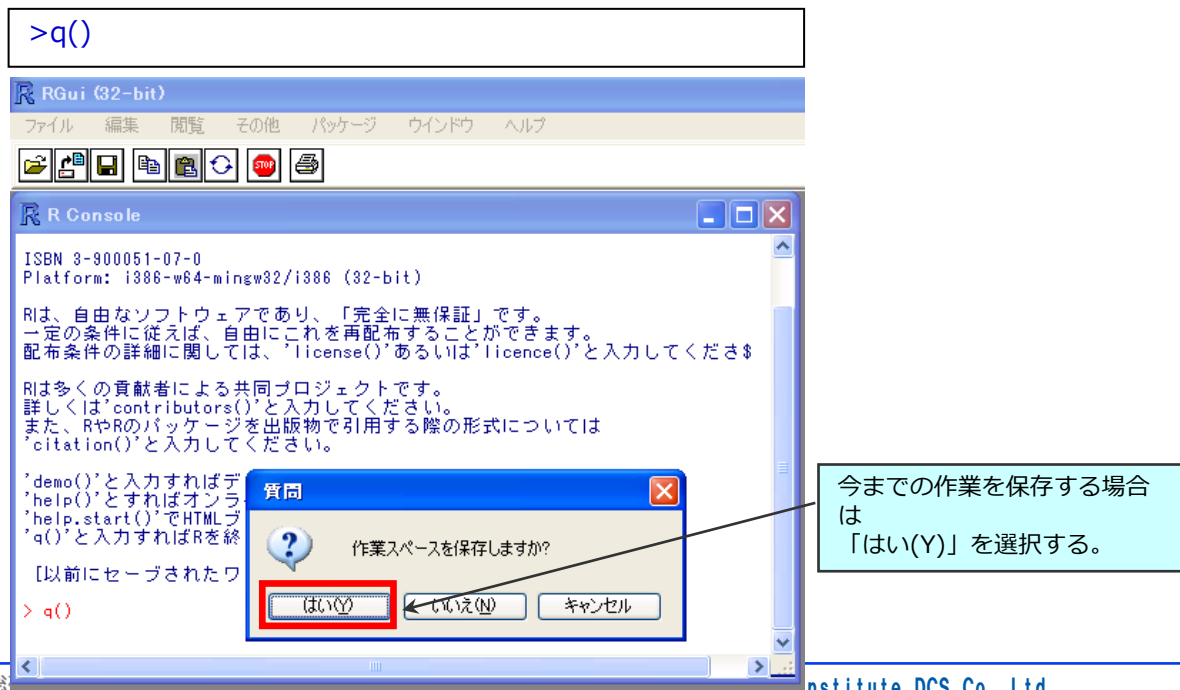
出力結果

```
data.frame*: 447 obs. of 42 variables:
 $ REAGUE      : chr "セ・リーグ" "セ・リーグ" "セ・リーグ" "セ・リーグ" ...
 $ TEAM        : chr "広島" "広島" "広島" "広島" ...
 $ Team_Order  : int 1 1 1 1 1 1 1 1 1 ...
 $ No          : int 0 0 2 4 5 6 7 9 10 13 ...
 $ PLAYER      : chr "中東 直己" "上本 崇司" "田中 広輔" "小窪 哲也" ...
 $ POSITION     : chr "外野手" "内野手" "内野手" "内野手" ...
 $ BIRTHDAY    : chr "1981/10/5" "1990/8/22" "1989/7/3" "1985/4/12" ...
 $ AGE         : int 35 26 27 31 36 36 25 27 30 33 ...
 $ ZAISEKI_NEN : int 10 4 3 9 4 11 7 9 8 1 ...
 $ HEIGHT      : int 168 170 171 175 187 173 183 177 182 185 ...
 $ WEIGHT      : int 73 71 81 80 99 76 83 90 96 93 ...
 $ BLOOD_TYPE  : chr "B型" "A型" "A型" "O型" ...
 $ TODA        : chr "右左" "右両" "右左" "右右" ...
 $ HOMETOWN    : chr "広島" "広島" "神奈川" "奈良" ...
 $ NEN_2016    : int 1300 880 4100 4000 12000 9000 1830 8500 1500 5700 ...
 $ NEN_2017    : num NA 700 7800 4000 NA 4000 1650 14000 1200 NA ...
 $ NEN_SAGAKU  : int NA -180 3700 0 NA -5000 -180 5500 -300 NA ...
 $ UP_RATIO    : num NA 0.795 1.902 1 NA ...
 $ KAIJYO      : int 1 0 0 0 1 0 0 0 0 1 ...
 $ DARITSU     : num NA 0 0.265 0.217 0.272 0 0.25 0.291 0.333 NA ...
 $ SHIRAI      : int NA 7 143 69 67 7 47 143 27 NA ...
 $ DASEKI      : int NA 0 679 107 268 12 60 652 22 NA ...
 $ DASUU       : int NA 0 581 92 243 9 56 557 21 NA ...
 $ TOKUTEN     : int NA 0 102 8 35 1 6 98 2 NA ...
 $ ANDA        : int NA 0 154 20 66 0 14 162 7 NA ...
 $ TWEEBASE    : int NA 0 17 9 8 0 0 30 2 NA ...
 $ THREEBASE   : int NA 0 3 1 1 0 0 8 1 NA ...
 $ HR          : int NA 0 13 2 5 0 2 20 0 NA ...
 $ HIT         : int NA 0 216 37 91 0 20 268 11 NA ...
 $ DATEN       : int NA 0 39 10 34 0 2 90 5 NA ...
 $ TOURUI      : int NA 0 28 0 6 0 3 23 0 NA ...
 $ TOURUI_SOSHI : int NA 0 19 0 0 0 1 9 0 NA ...
 $ GIDA        : int NA 0 3 1 0 0 1 1 0 NA ...
 $ GISEIFURAI  : int NA 0 1 0 2 0 0 3 0 NA ...
 $ FOUR_BALL   : int NA 0 77 11 22 3 3 84 1 NA ...
 $ KEIEN       : int NA 0 1 0 0 0 0 1 0 NA ...
 $ DEAD_BALL   : int NA 0 17 3 1 0 0 7 0 NA ...
 $ SANSHIN     : int NA 0 119 23 47 0 11 107 8 NA ...
 $ DOUBLE_PLAY : int NA 0 1 2 5 2 0 9 0 NA ...
 $ SLG         : num NA 0 0.372 0.402 0.374 0 0.357 0.481 0.524 NA ...
 $ OPS         : num NA 0 0.367 0.321 0.332 0.25 0.288 0.389 0.364 NA ...
 $ zougaku_flg : int 0 0 1 0 0 0 0 1 0 0 ...
```

9. R画面の終了

Rを終了する際、右上の「x」ボタンをクリックして終了するか、入力画面上から以下のコマンドで終了することができます。

この時、「作業スペースを保存しますか？」と尋ねられるので、今までの作業を保存する場合は「はい(Y)」を選択してください。



10.R利用の基礎（1/3）

Rの基本的な使い方を紹介します。

- プロンプト“>”の後に関数や式を入力してリターンキーを押すと、結果や答えが下に表示されます。 例) > 1+1 と入力してリターンを押してください。
- Rでは空白は無視されるので、入れても入れなくてもOKです。（1+1でも1 + 1でも同じ）
- 式の入力途中でリターンを押したときには、下に+と表示されて入力途中になります。式の入力を途中で止めるときは、Escを押します。
- Rでは #の後ろは無視されるので、コメントを書くときなどは#の後に書きます。
- “変数 = 式”とすれば、式が変数に代入されます。

例) X=1+2 としてXを表示させてみる

- 代入の式を丸カッコで囲むと、代入と表示を同時に行います。

例) (x=1+2) を実行してみる

- Rは大文字と小文字を区別するので、“x”と“X”は別の変数として扱われます。

10.R利用の基礎（2/3）

Rの基本的な使い方を紹介します。

- データフレームの基本操作は以下です。

列番号を指定して取得	<code>base[,2]</code>
複数の列番号を指定して取得	<code>base[,c(1,3)]</code>
列名で取得	<code>base[, "選手名"]</code>
<code>\$</code> 列名 でも取得できる	<code>base\$選手名</code>
複数の列名で取得	<code>base[,c("選手名","本塁打")]</code>
20列目から41列目を残す方法	<code>base2=data.frame(base)[c(20:41)]</code>
3列目の変数を取り除く	<code>base2=base[,-3]</code>
	<code>base2=base</code> <code>base2\$球団順位2012年 = NULL</code>
条件に合った行だけ取り出す	<code>base2=base[base\$球団=="巨人",]</code>
	<code>base2=subset(base,球団=="巨人",)</code>
昇順でソート	<code>base[order(base\$三振),]</code>
降順でソート	<code>base[order(base\$出塁率 , decreasing=T),]</code>

10.R利用の基礎（3/3）

Rの基本的な使い方を紹介します。

- 質的変数（カテゴリー変数）の要約（集計）を行うには、関数table()を使います。

例) > table(base\$球団)

- 対象データの数値の要約をするには、関数summary()を使います。

例) > summary(base)

- 結果を指数表示させたくない場合は、以下のオプションを実行しておきます。

> options(scipen=100)

- 解析を実施する際に、欠損値があるとエラーになって結果が出力されない場合があります。
初めからデータに一つでも欠損のあるレコードを除外しておくには、関数na.omit()を使います。

例) > base2=na.omit(base)

- データフレームをcsvファイルに保存するには、関数write.csv()を使います。

例) > write.csv(base2, "base2.csv", quote=FALSE, row.names=FALSE)
quote=FALSEで値を""で囲わない、row.names=FALSEで行番号を出さない
パスを指定せずにファイル名を描いた場合は、Rの作業フォルダに出力

11.注意事項

Rは大文字と小文字を区別します。
変数を指定する際は、正しく入力するよう注意が必要です。

※ Rの基本操作方法は、研修受講前に確認をお願いします。

12.（お願い）研修資料の取扱いについて

本研修で使用する研修テキスト及びサンプルデータの取扱いは、以下の通り注意して頂きますようお願い致します。

- 本研修テキストの一部、全体を問わず、承諾なく引用複製することを禁じます。
- テキスト及びサンプルデータは、本研修以外の目的で他者へコピーし提供することを禁じます。