

# Rstudio

データ分析やプログラミングを行うのに便利な機能が豊富

## Source

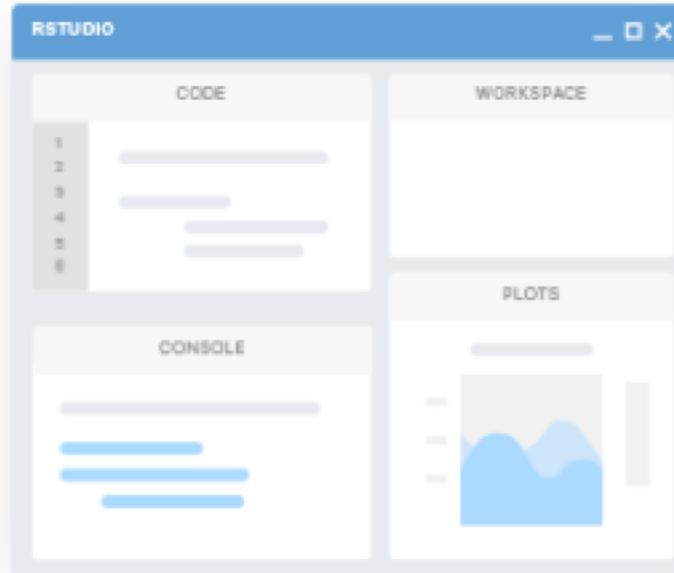
Rのコードやコメント  
が書かれたソース  
コードを表示

## Console

Sourceペインに書かれ  
たコードを実行した  
り、コマンドを直接  
入力して実行

## 補完機能

コードを入力している途中でコマンドの候補を提案してくれる  
長いパッケージ名、関数名やスペーリングを忘れてしまった時に便利  
入力途中で**Tabキーを打つ**ことでもコード補完機能を利用できる  
作成したオブジェクトの要素、ファイル名に対しても有効



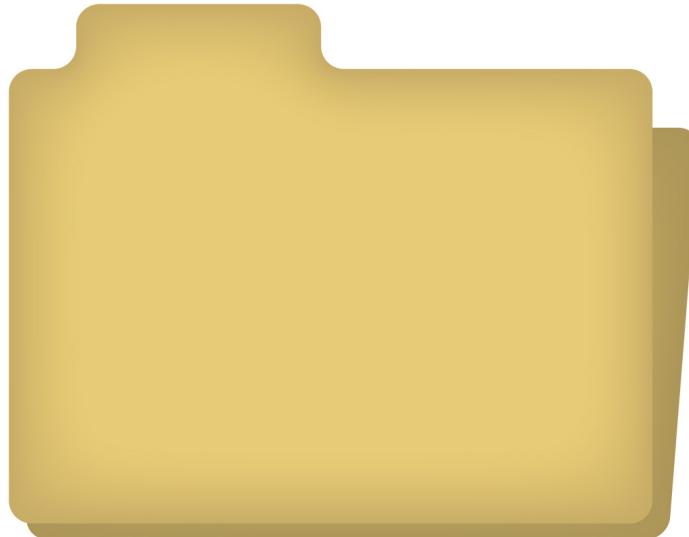
Environment ペインに読み込んだデー  
タや作成した変数の一覧を表示  
データを確認しながら進められる



# 作業ディレクトリとは？

作業ディレクトリとはデータを操作するフォルダのことです。

フォルダの中から必要なファイルを読み込んだり、結果を書き出して保存します。



例えば、同じフォルダの中に  
geno\_bc.csv  
rqtl.R (rqtl.Rmd)  
yield.csv  
を保存しておく

作業するのに必要なデータが保存してあるフォルダ

現在の作業ディレクトリがどこになっているのかを知るために

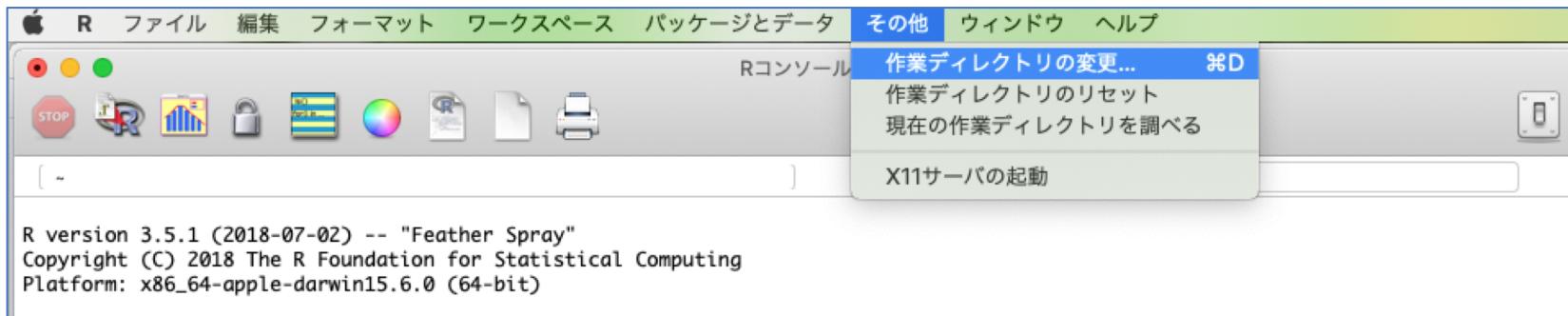
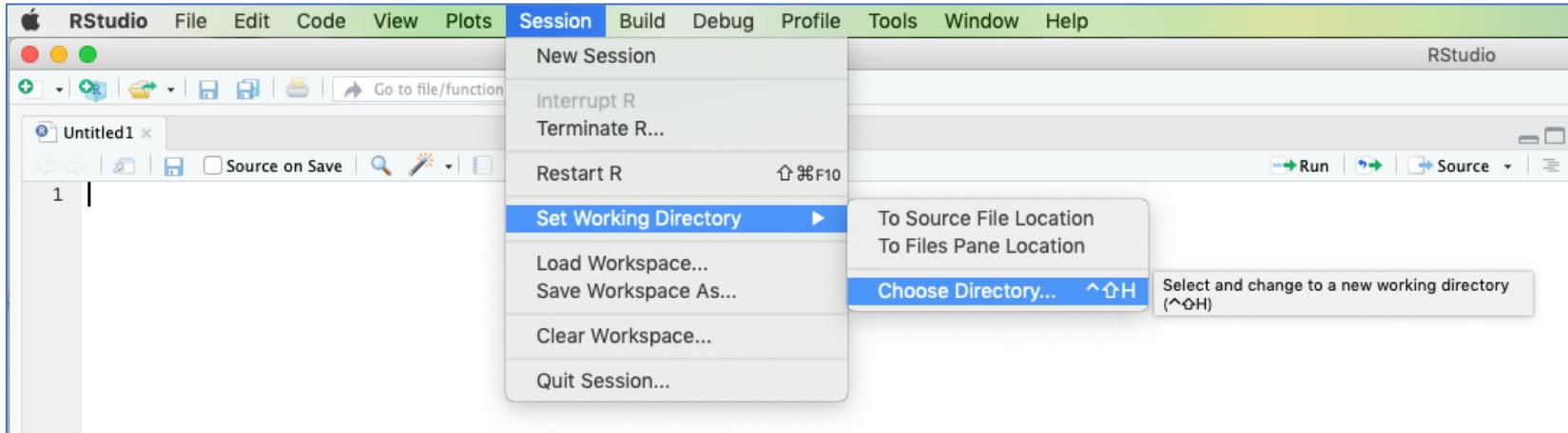
```
getwd()
```

と入力すると文字列で表示されます。

# 作業ディレクトリの変更

コマンドラインから、新しい作業ディレクトリを選択する `setwd("C:/Users/ユーザー名/Documents")`

マウス操作で行う方法



Windows の場合、メニューの [ファイル] から [ディレクトリの変更...] を選択してから、ディレクトリを変更することが出来る。

# パッケージのインストール

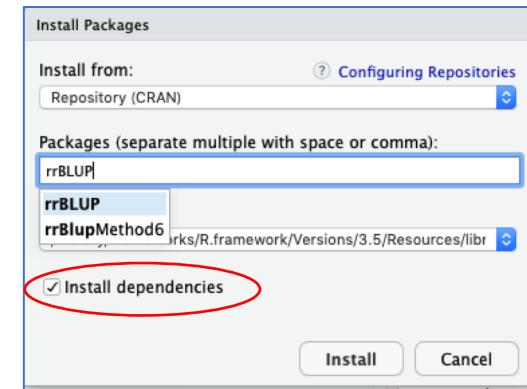
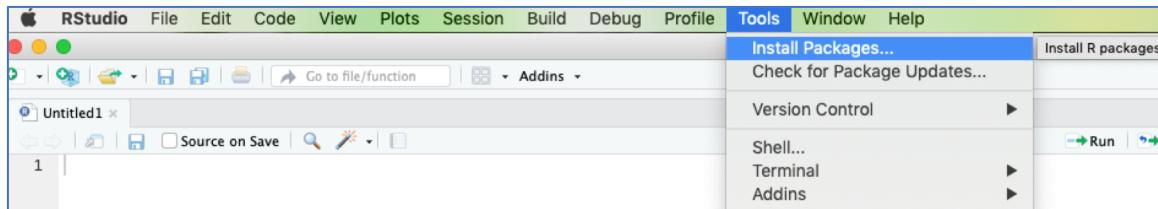
install.packages にパッケージの名前を指定

dependencies=TRUE を指定すると依存関係のあるパッケージもまとめてインストール

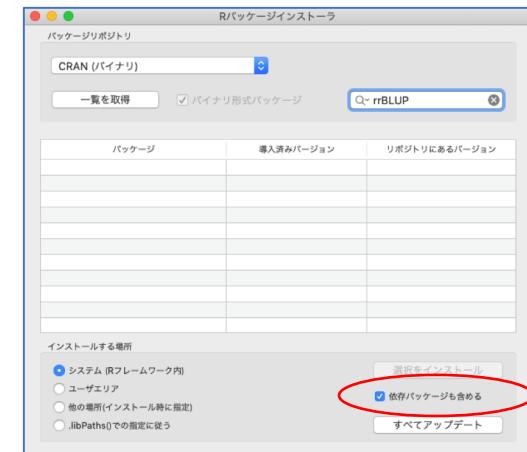
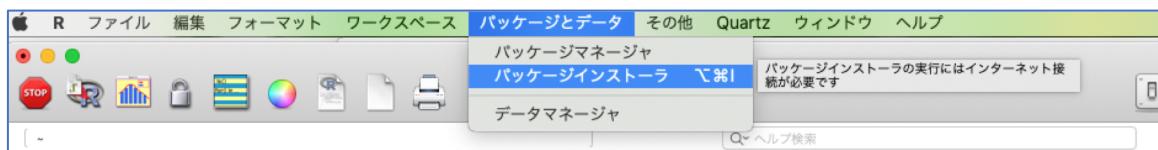
```
install.packages("rrBLUP", dependencies = TRUE)
```

実行すると、ミラーサイトの場所が聞かれるので、もっとも近い場所を選ぶ

マウス操作で行う方法



rrBLUP  
BGLR  
qtl



# R-Tips

- コマンドの履歴をさかのぼる
  - コンソールで上矢印のキーを押すと、以前に入力したコマンド履歴をさかのぼることができる

```
> str(iris)
'data.frame': 150 obs. of 5 variables:
 $ Sepal.Length: num 5.1 4.9 4.7 4.6 5 5.4 4.6 5 4.4 4.9 ...
 $ Sepal.Width : num 3.5 3 3.2 3.1 3.6 3.9 3.4 3.4 2.9 3.1 ...
 $ Petal.Length: num 1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ Petal.Width : num 0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ Species      : Factor w/ 3 levels "setosa","versicolor",...: 1 1 1 1 1 1 1 1 1 1 ...
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1        3.5         1.4        0.2   setosa
2          4.9        3.0         1.4        0.2   setosa
3          4.7        3.2         1.3        0.2   setosa
4          4.6        3.1         1.5        0.2   setosa
5          5.0        3.6         1.4        0.2   setosa
6          5.4        3.9         1.7        0.4   setosa
> head(iris)
```

上矢印や下矢印のキーを押すと、以前のコードが表示される  
ここでEnterキーを押せば、再度同じ計算を実行してくれる

# Rスクリプトを使ったデータ解析の手順

## アドホックなデータ解析

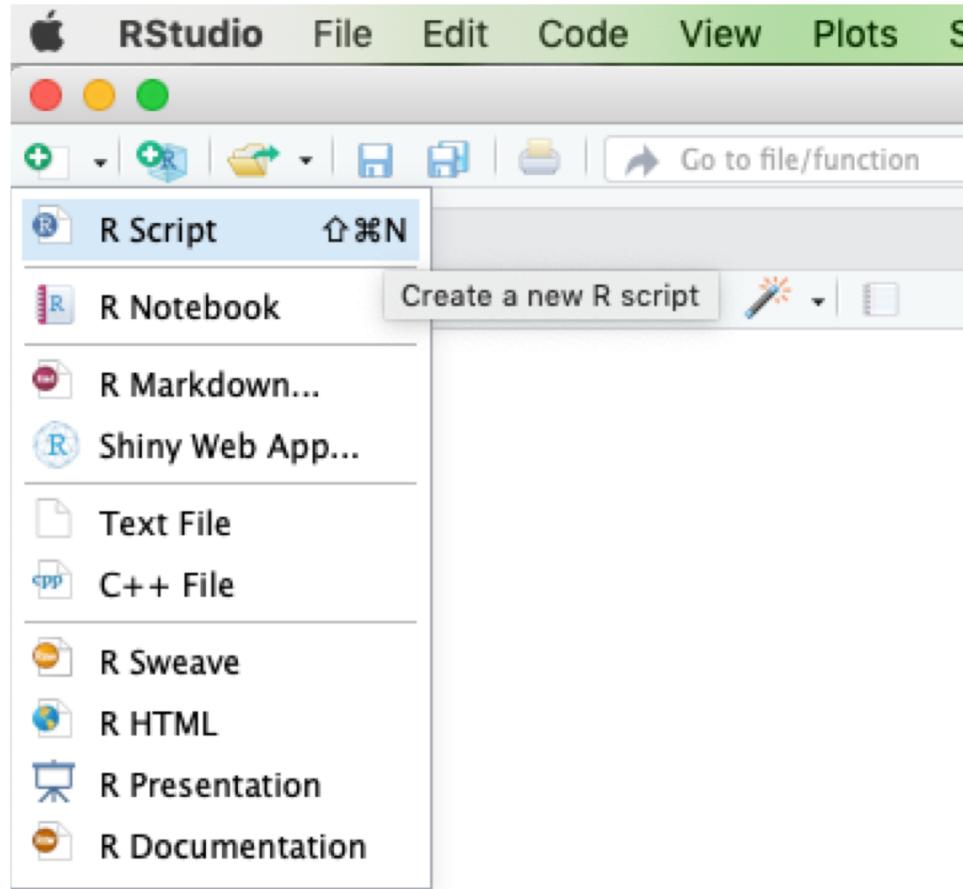
1. 実行するコマンドをコンソールに入力して実行
2. 結果を確認する
3. 欲しい結果が得られるまで、コマンドを修正  
(Rスクリプトに記録するというステップが抜けている)

## 再現可能なデータ解析

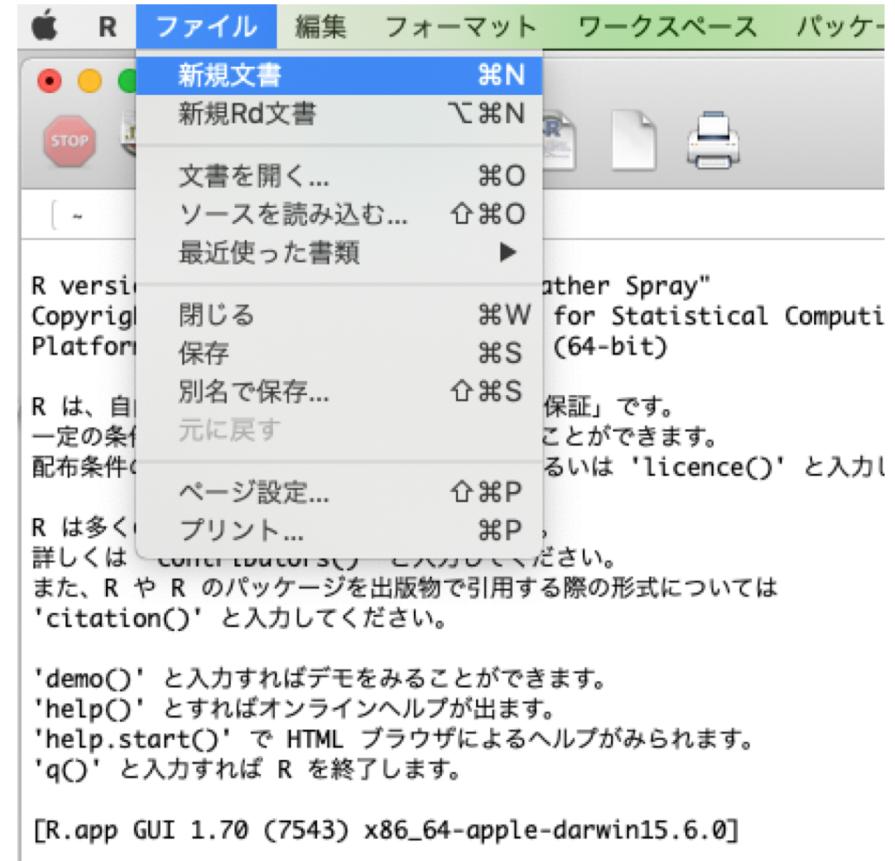
1. 必要な処理を実行するコマンドをRスクリプトに記述
2. 必要な箇所のコードを実行する
3. 結果を確認する
4. 欲しい結果が得られるまで、Rスクリプト上でコマンドの修正・実行・確認を繰り返す

# Rスクリプトの新規作成

## RStudio



## R



ファイル名を決めて保存

ファイル名に日本語や全角文字、スペースや特殊な記号を使うのは、あとで困ることになる可能性があるので避ける

# 「新しいスクリプト」でのコードの実行

## RStudio



The screenshot shows the RStudio interface with the title bar "RStudio". A toolbar at the top has several icons, including a "Run" icon with a green arrow and a red circle, which is highlighted with a red oval. Below the toolbar is a code editor window titled "190110GWAS.R". The code is written in R, starting with comments about the script's purpose and date, followed by instructions to set the working directory and read a CSV file named "gwas\_genotype.csv". The line "geno <- read.csv("gwas\_genotype.csv", row.names = 1)" is highlighted with a blue selection bar. To the right of the code editor, there is explanatory text in Japanese.

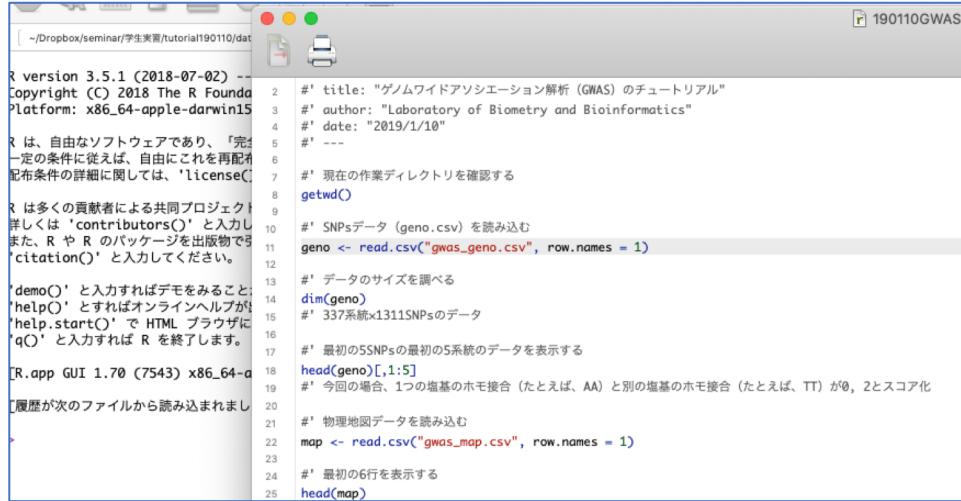
```
1 #' ---
2 #' title: "ゲノムワイドアソシエーション解析 (GWAS) のチュートリアル"
3 #' author: "Laboratory of Biometry and Bioinformatics"
4 #' date: "2019/1/10"
5 #' ---
6
7 #' 現在の作業ディレクトリを確認する
8 getwd()
9
10 #' SNPsデータ (geno.csv) を読み込む
11 geno <- read.csv("gwas_genotype.csv", row.names = 1)
12
13 #' データのサイズを調べる
14 dim(geno)
15 #' 337系統×1311SNPsのデータ
16
```

コードのある行を実行すると、RStudioは自動的にカーソルを次の行に送ります。

- Windowsの場合
  - ✓ 1行だけ実行したいときは、カーソルがその行にある状態で "control + R" を使います。
  - ✓ 複数行のプログラムを実行する場合は、実行する行を全て選択してから "control + R" を使います。
- Macの場合
  - ✓ 1行だけ実行したいときは、カーソルがその行にある状態で "command + return" を使います。
  - ✓ 複数行のプログラムを実行する場合は、実行する行を全て選択してから "command + return" を使います。

# 「新しいスクリプト」でのコードの実行

R



```
R version 3.5.1 (2018-07-02) -->
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15

R は、自由なソフトウェアであり、「完全に開かれた条件に従えば、自由にこれを再配布する。」を含む「GPLv2+」のライセンスで配布されています。詳細については、'license()' を参照してください。

R は多くの貢献者による共同プロジェクトです。詳しくは 'contributors()' と入力します。また、R や R のパッケージを出版物で引用する際には、'citation()' と入力してください。

'demo()' と入力すればデモをみることができます。
'help()' とすればオンラインヘルプが表示されます。
'help.start()' で HTML ブラウザに
'q()' と入力すれば R を終了します。

[R.app GUI 1.70 (7543) x86_64-aarch64]

[履歴が次のファイルから読み込まれます]
> 
  2   #' title: "ゲノムワイドアソシエーション解析 (GWAS) のチュートリアル"
  3   #' author: "Laboratory of Biometry and Bioinformatics"
  4   #' date: "2019/1/10"
  5   #' ---
  6
  7   #' 現在の作業ディレクトリを確認する
  8   getwd()
  9
 10  #' SNPsデータ (geno.csv) を読み込む
 11  geno <- read.csv("gwas_genotype.csv", row.names = 1)
 12
 13  #' データのサイズを調べる
 14  dim(geno)
 15  #' 337系統x1311SNPsのデータ
 16
 17  #' 最初の5SNPsの最初の5系統のデータを表示する
 18  head(geno)[,1:5]
 19  #' 今回の場合、1つの塩基のホモ接合（たとえば、AA）と別の塩基のホモ接合（たとえば、TT）が0、2とスコア化
 20
 21  #' 物理地図データを読み込む
 22  map <- read.csv("gwas_map.csv", row.names = 1)
 23
 24  #' 最初の6行を表示する
 25  head(map)
```

- Windowsの場合

- ✓ 1行だけ実行したいときは、カーソルがその行にある状態で "control + R" を使います。
- ✓ 複数行のプログラムを実行する場合は、実行する行を全て選択してから "control + R" を使います。

- Macの場合

- ✓ 1行だけ実行したいときは、カーソルがその行にある状態で "command + return" を使います。
- ✓ 複数行のプログラムを実行する場合は、実行する行を全て選択してから "command + return" を使います。

commands.R に R のコマンドを記述してある場合、以下のようにして実行できます。

```
> source("commands.R")
```

# ソース文書内でテキストの検索置換

The screenshot shows the RStudio interface with a script file titled '190110GWAS.R'. The search bar at the top contains the word 'geno'. Below the search bar, there are several search options: 'In selection', 'Match case', 'Whole word', 'Regex', and 'Wrap', with 'Wrap' checked. The main code area displays the following R script:

```
1 #' ---
2 #' title: "ゲノムワイドアソシエーション解析 (GWAS) のチュートリアル"
3 #' author: "Laboratory of Biometry and Bioinformatics"
4 #' date: "2019/1/10"
5 #' ---
6
7 #' 現在の作業ディレクトリを確認する
8 getwd()
```

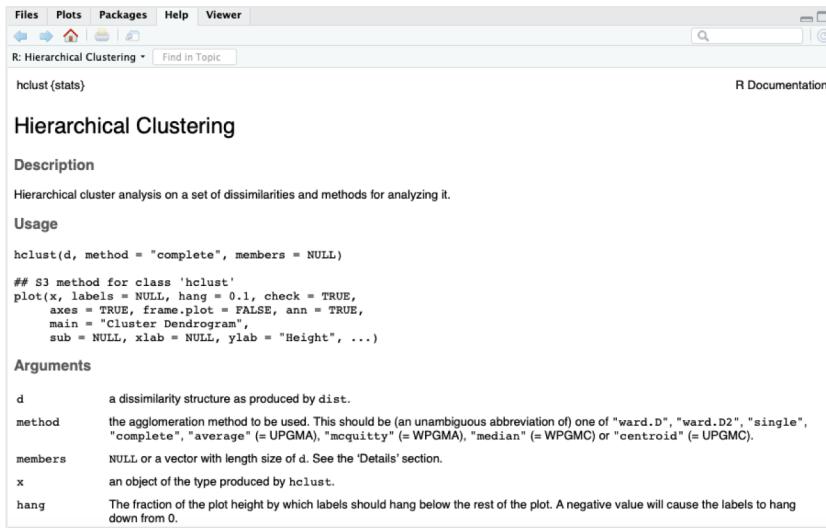
- RStudioはソース文書内でテキストの検索置換が出来ます
- 検索置換はCtrl-Fというショートカットキーで、あるいはEdit -> Find and Replaceというメニュー項目から開くことができます。

<http://memorandum2015.sakura.ne.jp/docs/using/source.html>

# 関数の利用方法を調べる

- `hclust`について調べるには、以下の2通りの方法でヘルプを見ることがある

```
help(hclust)  
? hclust
```



関数の完全な解説は必要無いが、引数のみを知りたいという場合は

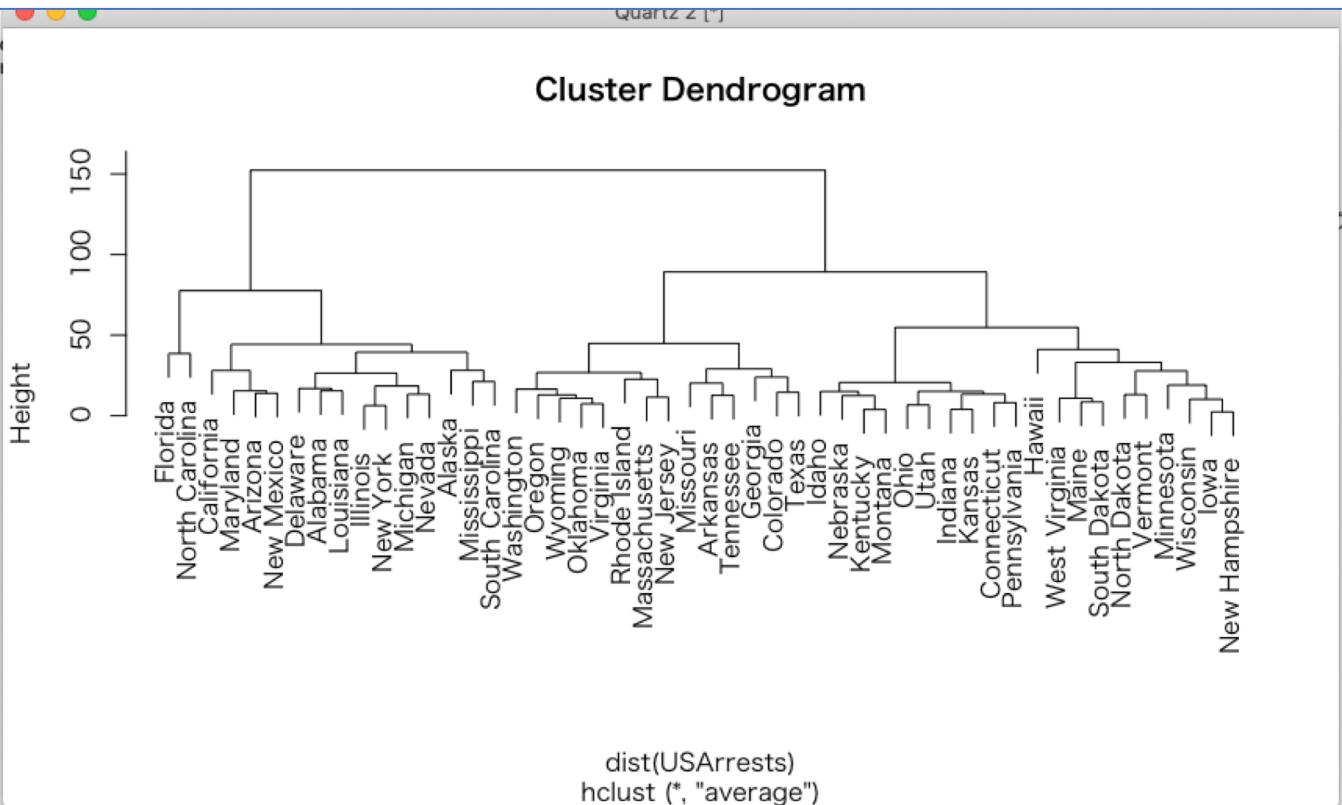
```
> args(hclust)  
function (d, method = "complete", members = NULL)  
NULL
```

# 関数の利用方法を調べる

関数のサンプルコードを実行

`example(hclust)`

```
[ワーカースペースが次のファイルから読み込まれました /Users/hkane/.RData  
[履歴が次のファイルから読み込まれました /Users/hkane/.Rapp.history  
> example(hclust)  
  
hclust> require(graphics)  
  
hclust> ### Example 1: Violent crime rates by US state  
hclust>  
hclust> hc <- hclust(dist(USArrests), "ave")  
  
hclust> plot(hc)  
次の図を見るためには <Return> キーを押して下さい:  
  
hclust> plot(hc, hang = -1)  
次の図を見るためには <Return> キーを押して下さい:
```



# Rを終了するには？

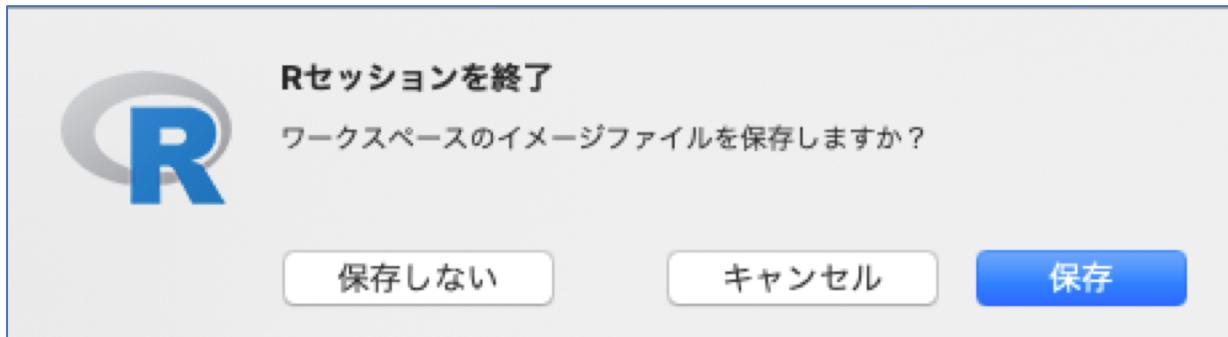
`q()`

`quit()`

と入力することで終了します。ウィンドウのXをクリックしても終了できます。

別ウィンドウで「作業スペースを保存しますか？」と聞かれますが、まずは保存しなくて大丈夫です。

次回以降も使うデータがたくさん場合など、また読み込むのは手間がかかる場合に、読み込んだデータ等を保存しておくためのものです。



# An Introduction To R

Rの操作でわからないことがあるとみるサイト 「The R Tips—データ解析環境Rの基本技・グラフィックス活用集」

<http://cse.naro.affrc.go.jp/takezawa/r-tips/r.html>

Rに関する情報交換を目的としたWiki 「RjpWiki」

<http://www.okadajp.org/RWiki/>

講義資料やサンプルコードも充実 「アグリバイオインフォマティクス教育研究ユニット」

<http://www.iu.a.u-tokyo.ac.jp/index.shtml>

Rで色を細かく指定するときに便利なサイト 「統計グラフの色」

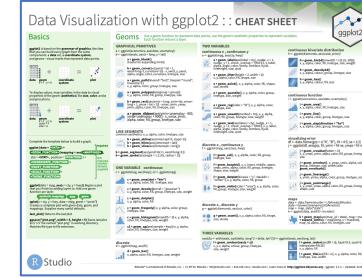
<https://oku.edu.mie-u.ac.jp/~okumura/stat/colors.html>

Rのパッケージを検索するサイト 「Microsoft R Application Network」

<https://mran.microsoft.com/packages>

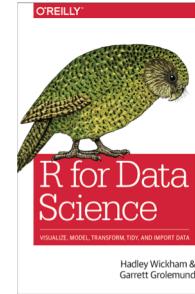
# Cheatsheets

<https://www.rstudio.com/resources/cheatsheets/>

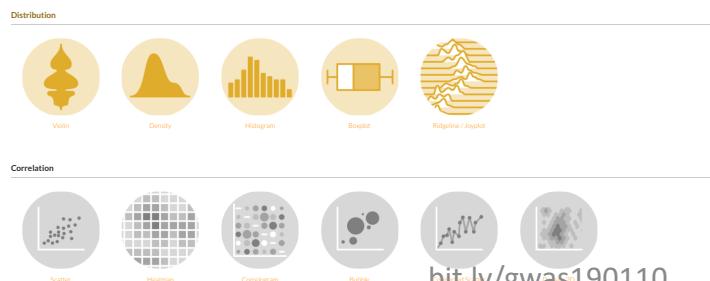


# R for Data Science

<https://r4ds.had.co.nz/>



# THE R GRAPH GALLAERY



# 参考図書

- 「みんなのR」  
出版社: マイナビ出版
- 「RユーザのためのRstudio[実践]入門」  
出版社: 技術評論社
- 「再現可能性のすゝめ (Wonderful R 3)」  
出版社: 共立出版
- 「Rではじめるデータサイエンス」  
出版社: オライリージャパン