



## Chapter02 補足資料

**RStudio Cloudを使用して、  
データ処理・解析を実行しよう**

# STEP1 RStudio Cloudとは？

- RStudioはプログラミング言語の一つであるR言語を扱いやすくした統合開発環境で、疫学分野では多くのユーザーによって利用されています。
- RStudioを使用するにはローカル環境を構築する必要がありますが、クラウド版のRStudio cloudではアカウントを作成するだけで、ローカル環境とほぼ同じことがクラウド上で実行できます。
- 誰もが無料で使えます。



・オフライン環境を構築する場合は以下のURLから  
<https://posit.co/download/rstudio-desktop/>

# STEP1 RStudio Cloudとは？

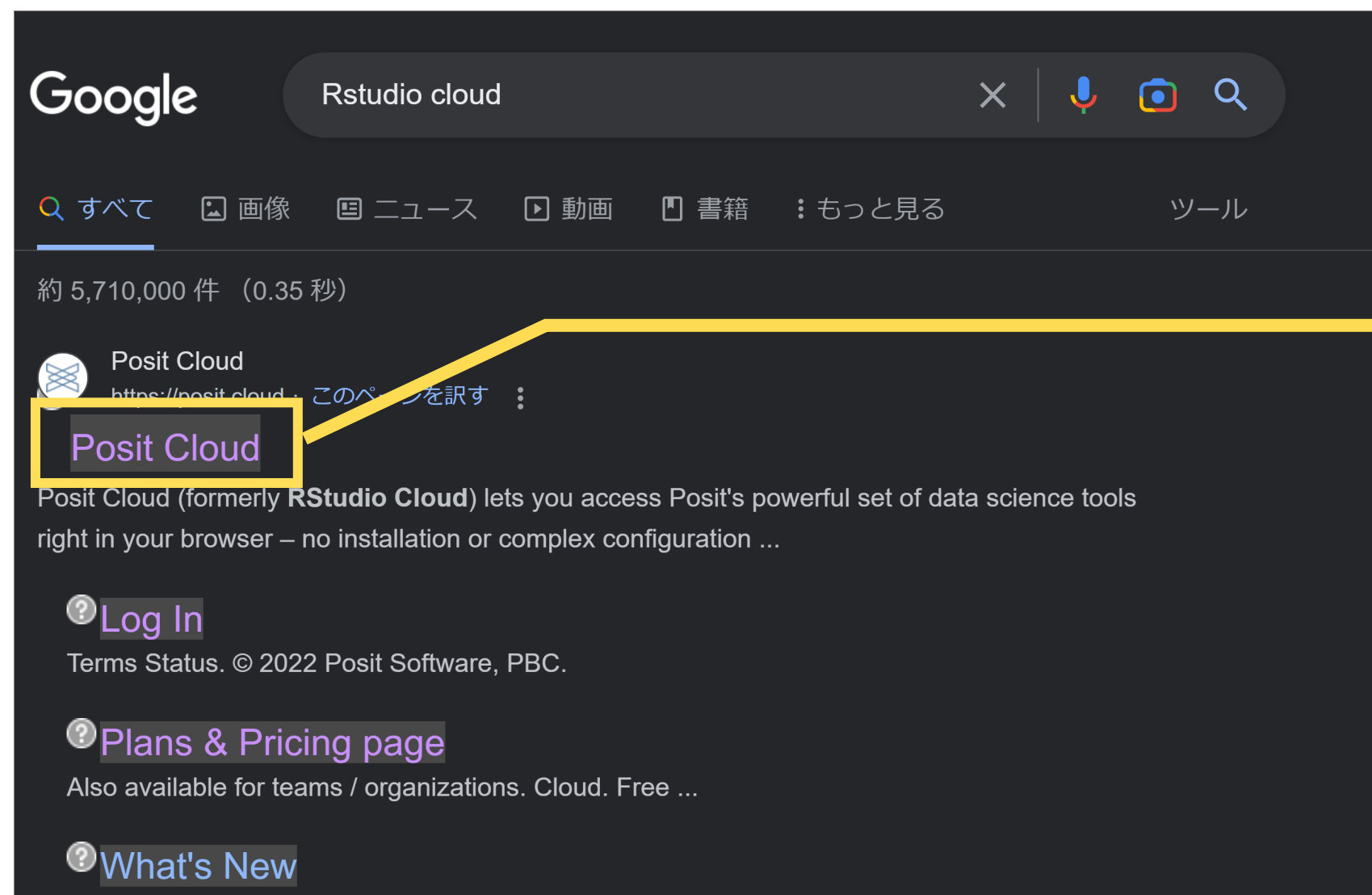
- 注意書き

この資料では、配布したRファイルのコードを実行して、Chapter02の内容を実施するための最低限の説明に留めています。全てのコードが予め記述されていますが、R言語の文法や構文などの説明は基本的にはありません。なお、R言語の学習については、スライドの最後におすすめ情報を掲載していますので、是非そちらをご確認ください。



・ オフライン環境を構築する場合は以下のURLから  
<https://posit.co/download/rstudio-desktop/>

# STEP2 RStudio Cloudの設定をしよう



Posit Cloudでアカウントを作成します。

1. 「RStudio Cloud」と検索します。
2. 「Posit Cloud」をクリックします。



## STEP2 RStudio Cloudの設定をしよう

 **posit** Cloud

Log In

Sign Up

# Friction free data science

Posit Cloud (formerly RStudio Cloud) lets you access Posit's powerful set of data science tools right in your browser – no installation or complex configuration required.

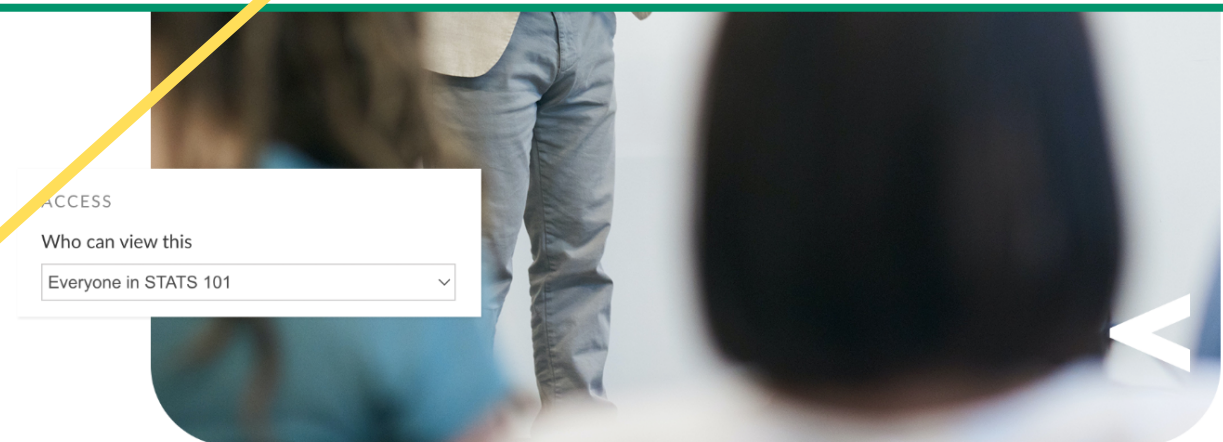
**GET STARTED**

**ALREADY A USER? LOG IN**

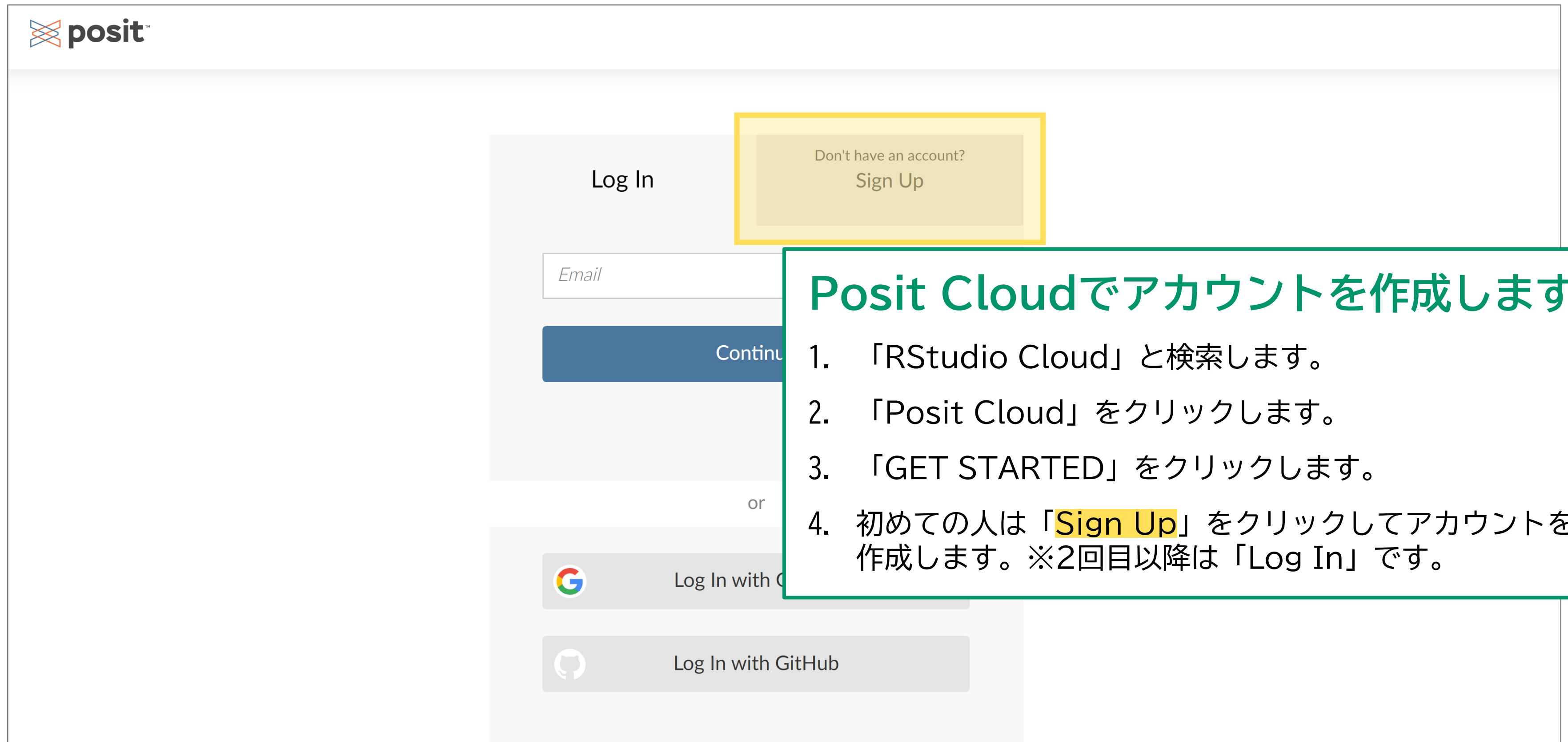
If you already have a shinyapps.io account, you can log in using your existing credentials.

## Posit Cloudでアカウントを作成します。

1. 「RStudio Cloud」と検索します。
2. 「Posit Cloud」をクリックします。
3. 「**GET STARTED**」をクリックします。



## STEP2 RStudio Cloudの設定をしよう



The screenshot shows the Posit Cloud login page. At the top left is the Posit logo. Below it, there's a 'Log In' section with a 'Don't have an account? Sign Up' link highlighted by a yellow box. Below the 'Log In' section is an email input field and a 'Continue' button. Below the 'Continue' button is an 'or' separator. Below the 'or' separator are two buttons: 'Log In with Google' and 'Log In with GitHub'.

### Posit Cloudでアカウントを作成します。

1. 「RStudio Cloud」と検索します。
2. 「Posit Cloud」をクリックします。
3. 「GET STARTED」をクリックします。
4. 初めての人は「**Sign Up**」をクリックしてアカウントを作成します。※2回目以降は「Log In」です。

## STEP2 RStudio Cloudの設定をしよう



Already have an account?  
Log In

Email

Password

First name

Last name

Sign Up

or



Sign Up with Google

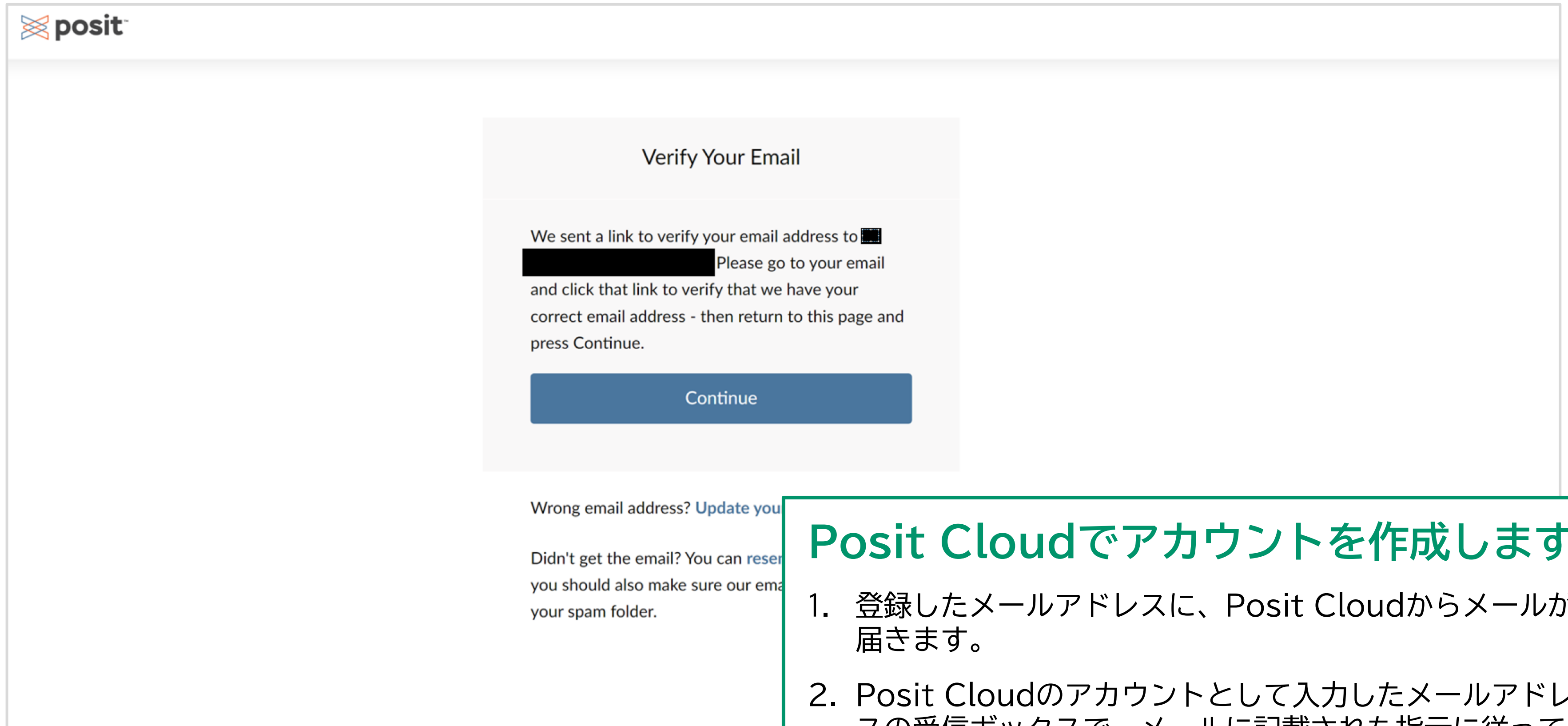



Sign Up with GitHub

### Posit Cloudでアカウントを作成します。



1. 「RStudio Cloud」と検索します。
2. 「Posit Cloud」をクリックします。
3. 「GET STARTED」をクリックします。
4. 初めての方は「Sign Up」をクリックしてアカウントを作成します。※2回目以降は「Log In」です。
5. 「Email」「Password」「First name」「Last name」を入力します。「Email」には、必ず使用できるアドレスを使ってください。  
※Googleアカウントを使用する場合は、「Sign Up with Google」をクリック

## STEP2 RStudio Cloudの設定をしよう





### Verify Your Email

We sent a link to verify your email address to   
 Please go to your email and click that link to verify that we have your correct email address - then return to this page and press Continue.

[Continue](#)

Wrong email address? [Update your email address](#)

Didn't get the email? You can [resend the email](#). You should also make sure our email is in your spam folder.

## Posit Cloudでアカウントを作成します。

1. 登録したメールアドレスに、Posit Cloudからメールが届きます。
2. Posit Cloudのアカウントとして入力したメールアドレスの受信ボックスで、メールに記載された指示に従って認証手続きを行ってください。

## STEP2 RStudio Cloudの設定をしよう

× 閉じる | 戻る 次へ

Please verify your email address



Posit Cloud <noreply@posit.cloud>  
宛先: 自分

To verify your email address, click on the link below:

? [Verify your email](#)

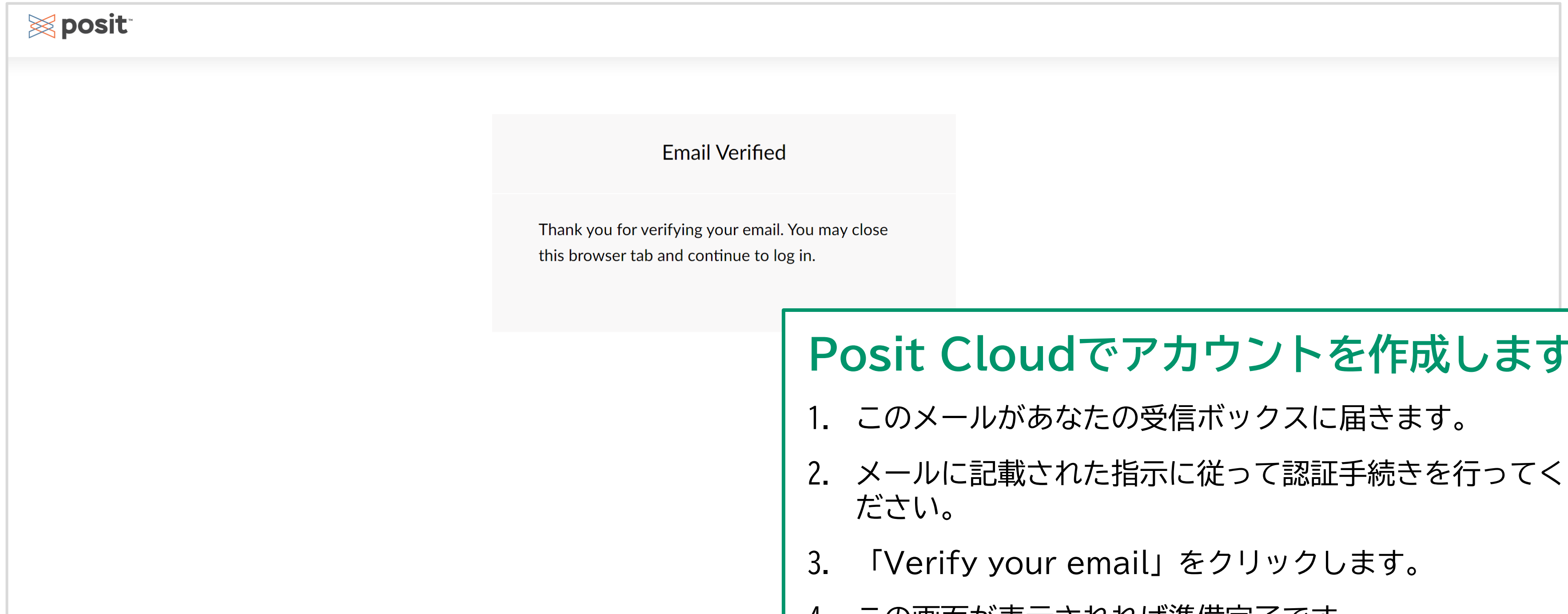
Please note this link will only be valid for 72 hours from the time this email was sent.

Thanks for using Posit Cloud!  
- The Posit Cloud Team

### Posit Cloudでアカウントを作成します。

1. このメールがあなたの受信ボックスに届きます。
2. メールに記載された指示に従って認証手続きを行ってください。
3. 「Verify your email」をクリックします。

## STEP2 RStudio Cloudの設定をしよう



### Posit Cloudでアカウントを作成します。

1. このメールがあなたの受信ボックスに届きます。
2. メールに記載された指示に従って認証手続きを行ってください。
3. 「Verify your email」をクリックします。
4. この画面が表示されれば準備完了です。  
Posit Cloudの画面に戻って、ログインしましょう。

## STEP2 RStudio Cloudの設定をしよう



### Posit Cloudでアカウントを作成します。

1. このメールがあなたの受信ボックスに届きます。
2. メールに記載された指示に従って認証手続きを行ってください。
3. 「Verify your email」をクリックします。
4. この画面が表示されれば準備完了です。  
Posit Cloudの画面に戻って、ログインしましょう。
5. この画面があなたのワークスペースとなります。

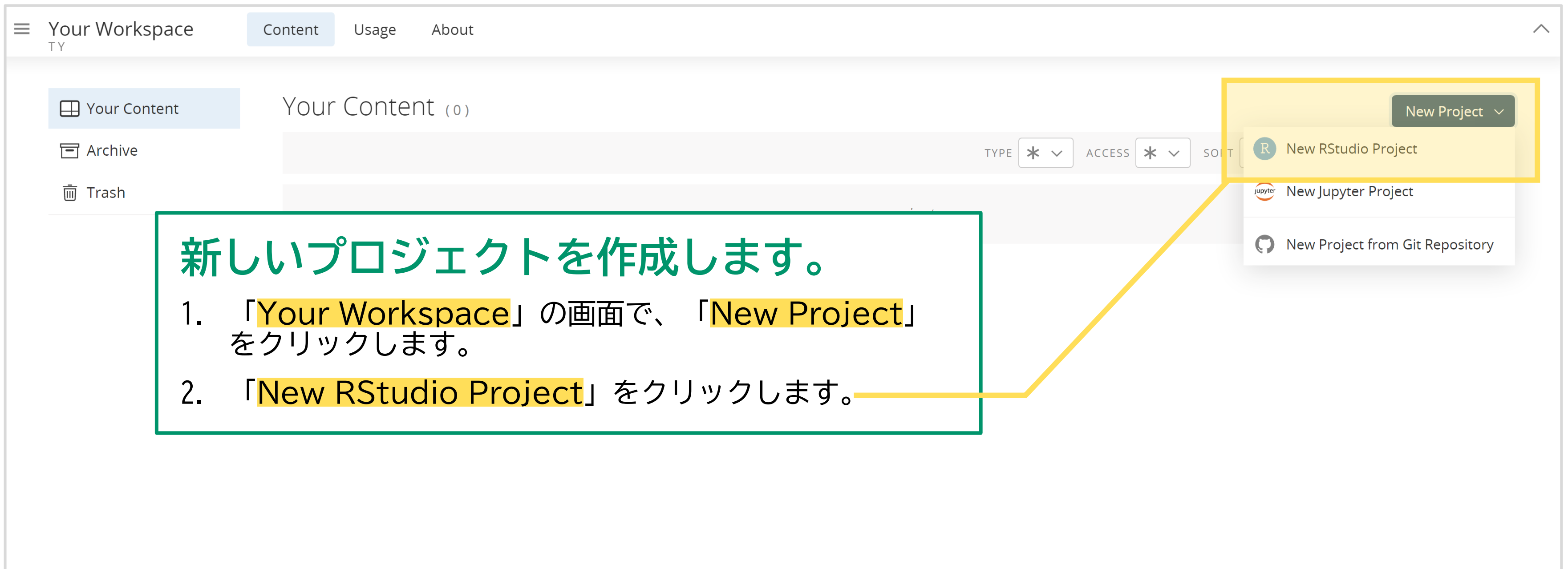
\*RStudio Cloudには無料使用の制限があります。詳しくは「project hours」という値で決まり、2023年3月25日現在では、「25 project hours/月」となっています。授業等での使用であれば、この時間を超えることはほぼないと思います。仮に、5ドル支払えば、75 project hour/monthまで利用することが可能となります。  
(利用時間が増える場合には、RStudioのローカル環境を構築することをお勧めいたします。)

## STEP3 Rファイルをアップロードしよう

- 配布されている以下の資料をすべてRStudio Cloudへアップロードしよう。
  - ・ File2-1.xlsx
  - ・ File2-3.R
  - ・ File2-5.xlsx
  - ・ File2-6.xlsx
  - ・ File2-7.xlsx
- RStudioにおけるコードの実行方法を知ろう。



## STEP3 Rファイルをアップロードしよう

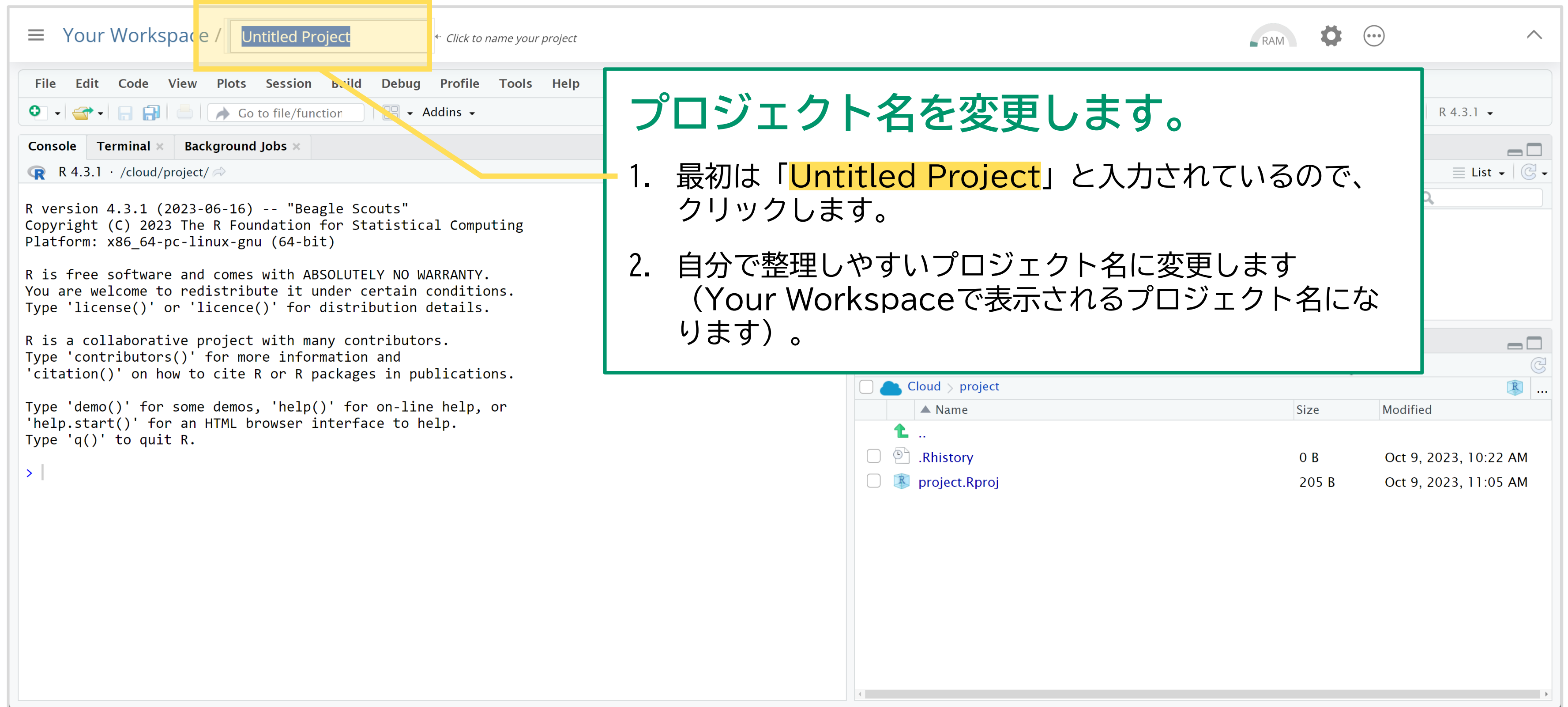


The screenshot shows the 'Your Workspace' interface with the 'Content' tab selected. On the left, there is a sidebar with 'Your Content', 'Archive', and 'Trash'. The main area is titled 'Your Content (0)' and has filters for 'TYPE', 'ACCESS', and 'SORT'. A 'New Project' button is in the top right, which has been clicked to reveal a dropdown menu. The menu options are 'New RStudio Project', 'New Jupyter Project', and 'New Project from Git Repository'. A yellow box highlights the 'New Project' button and the dropdown menu. A yellow line connects the 'New RStudio Project' option in the dropdown to the second step in the instructions below.

新しいプロジェクトを作成します。

1. 「Your Workspace」の画面で、「New Project」をクリックします。
2. 「New RStudio Project」をクリックします。

## STEP3 Rファイルをアップロードしよう



The screenshot shows the RStudio Cloud interface. At the top, the breadcrumb navigation bar displays 'Your Workspace / Untitled Project'. A yellow box highlights the 'Untitled Project' text, with a yellow arrow pointing to the first step of the instructions. The main console area on the left shows the R version 4.3.1 and its license information. The bottom right pane shows a file explorer for the 'project' directory, listing files like '.Rhistory' and 'project.Rproj'.

### プロジェクト名を変更します。

1. 最初は「**Untitled Project**」と入力されているので、クリックします。
2. 自分で整理しやすいプロジェクト名に変更します  
(Your Workspaceで表示されるプロジェクト名になります)。

Name	Size	Modified
..		
.Rhistory	0 B	Oct 9, 2023, 10:22 AM
project.Rproj	205 B	Oct 9, 2023, 11:05 AM

## STEP3 Rファイルをアップロードしよう

The screenshot shows the RStudio IDE interface. The top bar displays 'Your Workspace / Kenpakusha\_Rcode\_20230312'. The console window on the left shows the R version 4.3.1 startup message. A yellow box highlights the project name 'Kenpakusha\_Rcode\_20230312' in the top bar, with a yellow arrow pointing to the instruction box on the right.

### プロジェクト名を変更します。

1. 最初は「Untitled Project」と入力されているので、クリックします。
2. 自分で整理しやすいプロジェクト名に変更します  
(Your Workspaceで表示されるプロジェクト名になります)。
3. ここでは「**Kenpakusha\_Rcode\_20230312**」としました。

## STEP3 Rファイルをアップロードしよう

Upload Files

Target directory:  
/cloud/project

File to upload:  
ファイルを選択 選択されていません

TIP: To upload multiple files or a directory, create a zip file.  
The zip file will be automatically expanded after upload.

OK Cancel

Environment is empty

Environment

History

Connections

Tutorial

Import Dataset

161 MiB

R

Global Environment

packages

Help

New Blank File

Upload

Delete

Rename

More

	Size	Modified
0 B	Oct 9, 2023, 10:22 AM	
205 B	Oct 9, 2023, 2:14 PM	

project.Rproj

### Rファイルをアップロードします。

1. 右側の中段にある「Upload」をクリックします。
2. 中央に現れたダイアログにある「ファイルを選択」をクリックします。
3. PCに保存されている「File2-3.R」のファイルを選択します（該当のファイルは、教科書の指示に従って、ご自身のPCにダウンロードしてください）。

## STEP3 Rファイルをアップロードしよう

The screenshot shows the RStudio interface with the 'Upload Files' dialog box open. The dialog box has a 'Target directory' field set to '/cloud/project' and a 'File to upload' field containing 'File2-3.R'. The 'OK' button is highlighted. Red circles with numbers 4 and 5 indicate the steps to click 'File2-3.R' and 'OK' respectively.

Environment is empty

	Size	Modified
0 B	Oct 9, 2023, 10:22 AM	
205 B	Oct 9, 2023, 2:14 PM	

### Rファイルをアップロードします。

1. 右側の中段にある「Upload」をクリックします。
2. 中央に現れたダイアログにある「ファイルを選択」をクリックします。
3. PCに保存されている「File2-3.R」のファイルを選択します（該当のファイルは、教科書の指示に従って、ご自身のPCにダウンロードしてください）。
4. 上記3. で選択したファイルが表示されます。
5. 「OK」をクリックします。

## STEP3 Rファイルをアップロードしよう

Console

```
R 4.3.1 · /cloud/project/

R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

Environment

Environment is empty

Files

Name	Size	Modified
..		
.Rhistory	0 B	Oct 9, 2023, 10:22 AM
project.Rproj	205 B	Oct 9, 2023, 2:14 PM
File2-3.R	13.8 KB	Oct 9, 2023, 2:21 PM

アップロードされているか確認します。

1. 右下の枠に、「File2-3.R」があるか確認します。
2. ここに該当のファイルがあれば、クラウド上のフォルダへのアップロードが完了です。



## STEP3 Rファイルをアップロードしよう

≡ Your Workspace / Kenpakusha\_Rcode\_20230312

File Edit Code View Plots Session Build Debug Profile Tools Help

+ Go to file/function Addins

Console Terminal Background Jobs

R 4.3.1 · /cloud/project/

```
R version 4.3.1 (2023-06-16) -- "Beagle Scouts"
Copyright (C) 2023 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)
```

```
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
```

```
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
```

```
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
```

Environment History Connections Tutorial

Import Dataset 163 MiB

R Global Environment

Environment is empty

Files Plots Packages Help Viewer Presentation

New Folder New Blank File Upload Delete Rename More

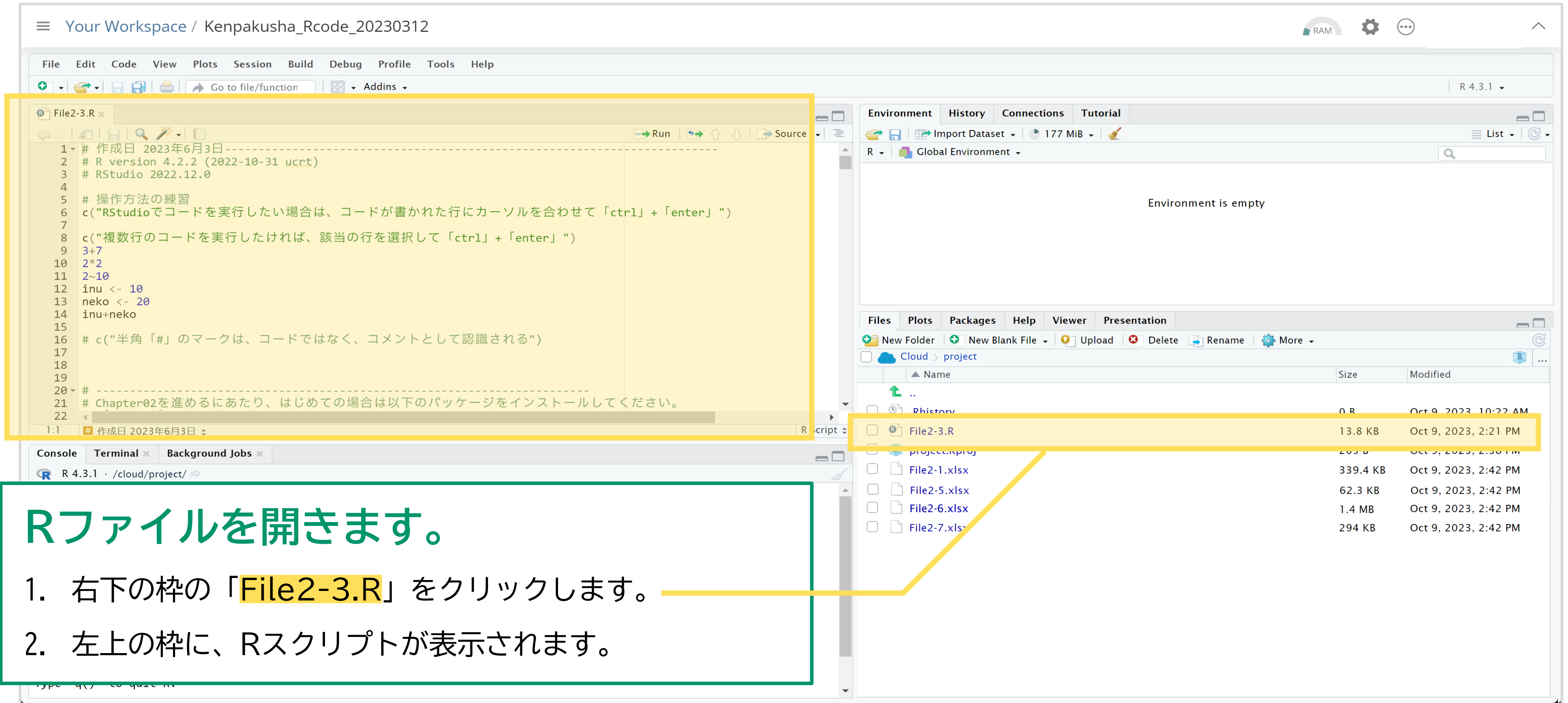
Cloud > project

	Name	Size	Modified
	..		
	File2-3.R	13.8 KB	Oct 9, 2023, 2:21 PM
	project.Rproj	205 B	Oct 9, 2023, 2:38 PM
	File2-1.xlsx	339.4 KB	Oct 9, 2023, 2:42 PM
	File2-5.xlsx	62.3 KB	Oct 9, 2023, 2:42 PM
	File2-6.xlsx	1.4 MB	Oct 9, 2023, 2:42 PM
	File2-7.xlsx	294 KB	Oct 9, 2023, 2:42 PM

### アップロードされているか確認します。

1. 右下の枠に、「File2-3.R」があるか確認します。
2. ここに該当のファイルがあれば、クラウド上のフォルダへのアップロードが完了です。
3. 同様の手順で、他のファイルのアップロードも完了させてください。
  - File2-1.xlsx
  - File2-5.xlsx
  - File2-6.xlsx
  - File2-7.xlsx

## STEP3 Rファイルをアップロードしよう



The screenshot shows the RStudio interface with the following components:

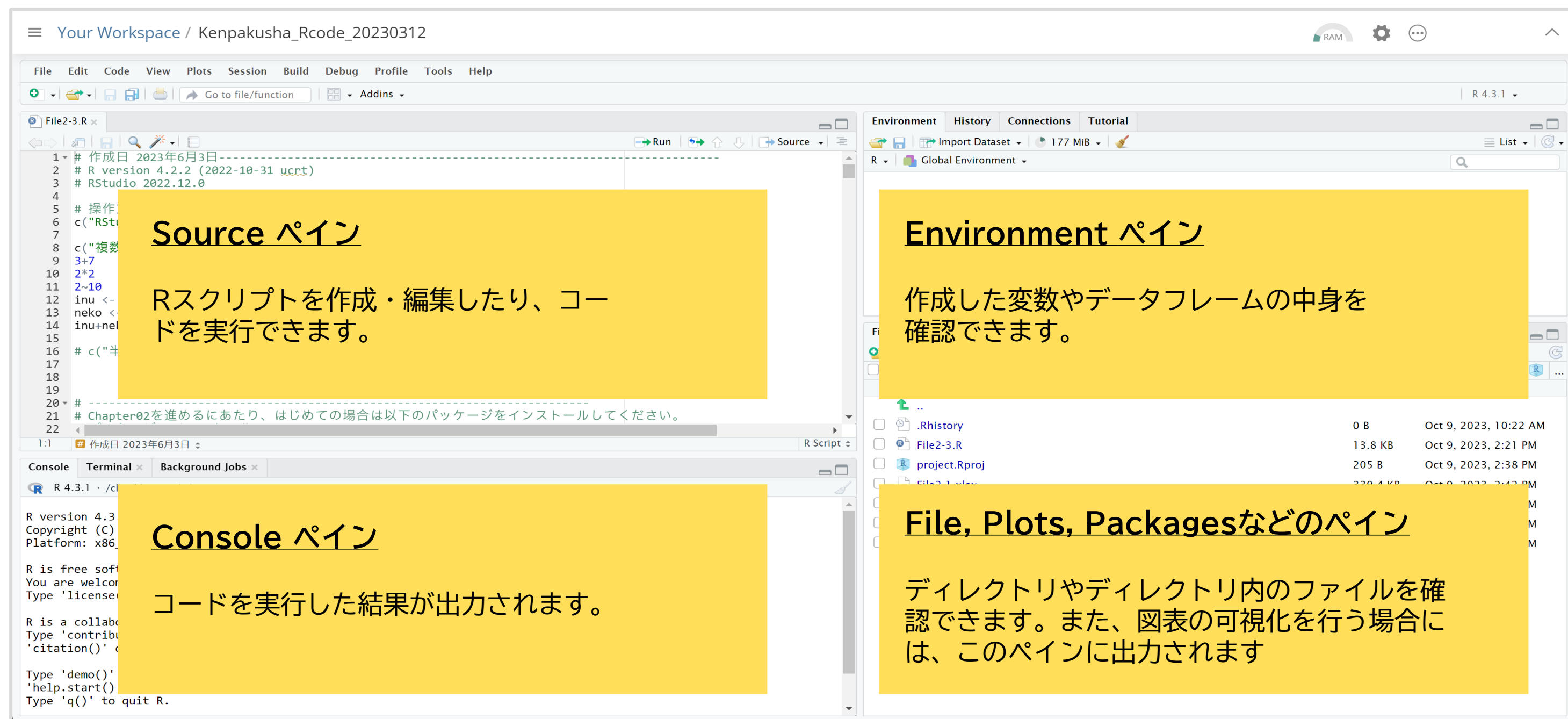
- Top Bar:** "Your Workspace / Kenpakusha\_Rcode\_20230312", RAM status, settings, and help icons.
- Menu Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Icons for file operations and a "Go to file/function" search bar.
- Left Pane (Source Editor):** Displays the content of "File2-3.R". The code includes comments in Japanese and R commands like `c("RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」 + 「enter」")`, `3+7`, `2*2`, `2~10`, `inu <- 10`, `neko <- 20`, and `inu+neko`. It also includes a comment about installing packages for Chapter 02.
- Right Pane (Environment/Files):** The "Files" tab is active, showing a list of files in the "project" directory. The file "File2-3.R" is highlighted.

**Rファイルを開きます。**

- 右下の枠の「File2-3.R」をクリックします。
- 左上の枠に、Rスクリプトが表示されます。



## STEP3 Rファイルをアップロードしよう



## Rファイルを開きます。

1. 4つの枠（パインと呼びます）にはそれぞれ名前があります。参考までに覚えておいてください。

## STEP3 Rファイルをアップロードしよう

スク립トにあるコードを実行してみよう。

- 基本操作として、Rでコードを実行する場合には以下の通りです。
  - Source ペインで、実行したいコードの行にカーソルを合わせます。
  - 「ctrl」 + 「enter」を押す。
- 試しに、6行目のコードを実行してみましょう。
- カーソルを6行目に移動させます。
- 6行目で「ctrl」 + 「enter」を押して、コードを実行します。
- 左下のコンソールに、実行結果が出力されます。

```
# 作成日 2023年6月3日-----
# R version 4.2.2 (2022-10-31 ucrt)
# RStudio 2022.12.0
# 操作方法の練習
6 c("RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」 + 「enter」")
8 c("複数行のコードを実行したければ、該当の行を選択して「ctrl」 + 「enter」")
9 3+7
10 2*2
11 2~10
12 inu <- 10
13 neko <- 20
14 inu+neko
15
16 # c("半角「#」のマークは、コードではなく、コメントとして認識される")
17
18
19
20 # -----
21 # Chapter02を進めるにあたり、はじめての場合は以下のパッケージをインストールして
22
```

```
R 4.3.1 · /cloud/project/
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> # 操作方法の練習
> c("RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」 + 「enter」")
[1] "RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」 + 「enter」"
```

## STEP3 Rファイルをアップロードしよう

スクリプトにあるコードを実行してみよう。

1. 一度に、複数行のコードを実行したい場合は、該当の行を選択します。
2. 「ctrl」 + 「enter」を押す。
3. 実行した行のすべての結果が、一度に出力されます。

≡ Your Workspace / Kenpakusha\_Rcode\_20230312

File Edit Code View Plots Session Build Debug Profile Tools Help

+ - Go to file/function Addins

File2-3.R x

```
1 # 作成日 2023年6月3日
2 # R version 4.2.2 (2022-10-31 ucrt)
3 # RStudio 2022.12.0
4
5 # 操作方法の練習
6 c("RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」+「enter」")
7
8 c("複数行のコードを実行したければ、該当の行を選択して「ctrl」+「enter」")
9 3+7
10 2*2
11 2~10
12 inu <- 10
13 neko <- 20
14 inu+neko
15
16 # c("半角「#」のマークは、コードではなく、コメントとして認識される")
17
18
19
20 # -----
21 # Chapter02を進めるにあたり、はじめての場合は以下のパッケージをインストールしてください。
22
```

Console Terminal x Background Jobs x

```
R 4.3.1 · /cloud/project/
> # 操作方法の練習
> c("RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」+「enter」")
[1] "RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」+「enter」"
> c("複数行のコードを実行したければ、該当の行を選択して「ctrl」+「enter」")
[1] "複数行のコードを実行したければ、該当の行を選択して「ctrl」+「enter」"
> 3+7
[1] 10
> 2*2
[1] 4
> 2~10
2 ~ 10
> inu <- 10
> neko <- 20
> inu+neko
[1] 30
>
```

Name	Size	Modified
..		
.Rhistory	0 B	Oct 9, 2023, 10:22 AM
File2-3.R	13.8 KB	Oct 9, 2023, 2:21 PM
project.Rproj	205 B	Oct 9, 2023, 2:38 PM
File2-1.xlsx	339.4 KB	Oct 9, 2023, 2:42 PM
File2-5.xlsx	62.3 KB	Oct 9, 2023, 2:42 PM
File2-6.xlsx	1.4 MB	Oct 9, 2023, 2:42 PM
File2-7.xlsx	294 KB	Oct 9, 2023, 2:42 PM

## STEP3 Rファイルをアップロードしよう

The screenshot shows the RStudio interface with a script file named 'File2-3.R' open. The script contains several lines of R code, including comments and assignments. A yellow box highlights a comment line: `# c("半角「#」のマークは、コードではなく、コメントとして認識される")`. The console at the bottom shows the output of the executed code, including the execution of the comment line, which results in a blank line. A red box highlights the prompt `>` in the console.

Environment History Connections Tutorial

R 4.3.1

Global Environment

Values

inu	10
neko	20

### スクリプトにあるコードを実行してみよう。

1. 行の先頭等に「#」のついている箇所は、コードとして認識されないようになっています。この行にカーソルを合わせて、「ctrl」＋「enter」で実行しても、#マークの行は飛ばされる（次の行のコードが実行される）ことになります。
2. なお、コードが正常に実行されると、コンソールのペインには左下の赤色部分のように、「>」の表示が最後にあります。

# STEP4 パッケージをインストールしよう

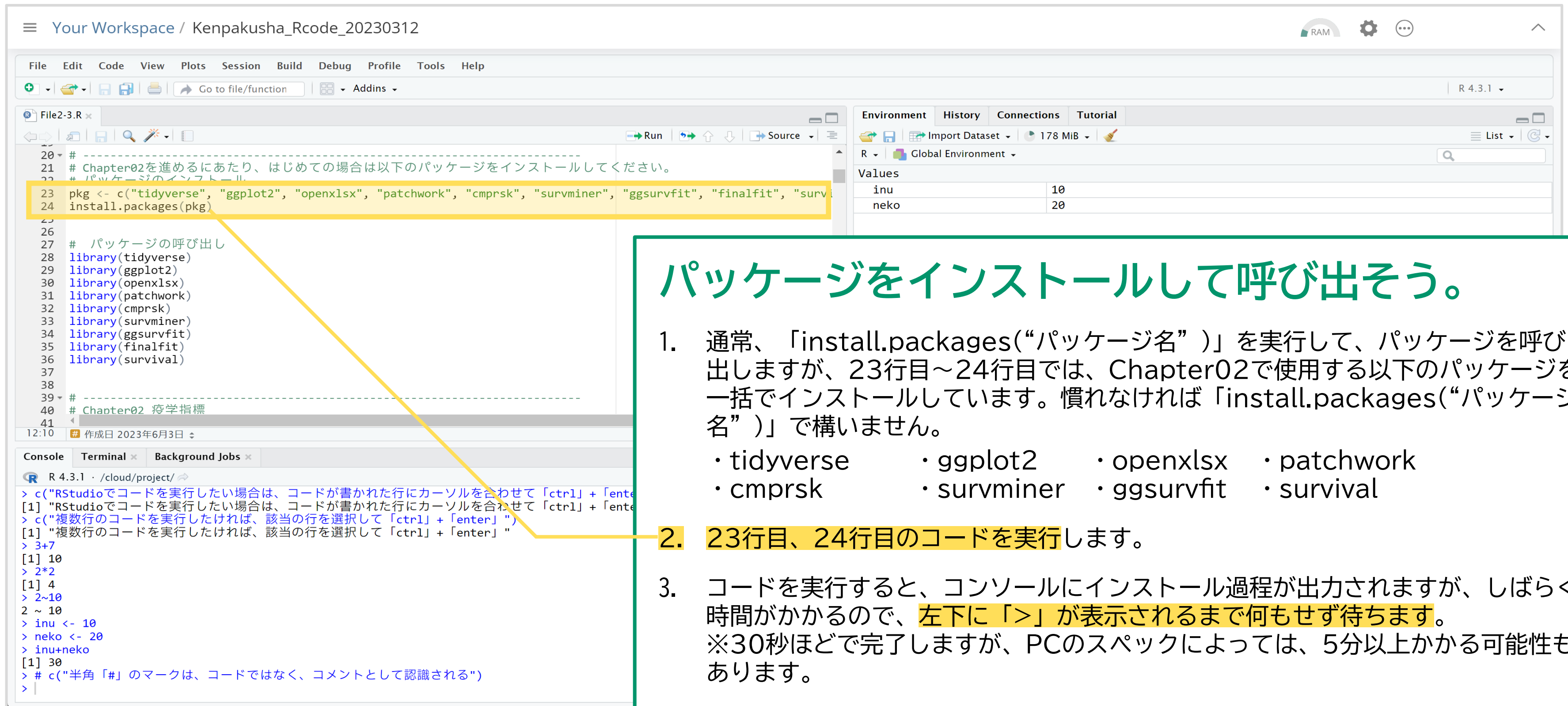
- R言語では、データ処理等を行うための基本パッケージが備わっていて、デフォルトの状態ですべての関数（sum, mean, plotなど）を使うことができます。
- さらに、便利機能を搭載した無数の“オプションパッケージ”が開発・公開されていて、これらを適宜使いながらデータ処理を行うことが多いです。
- オプションパッケージを使用した処理を行うためには、該当のパッケージを①インストールし、②呼び出す必要があります。
- PC環境が変わらなければ、インストールは初回のみでOKです。一方、パッケージを使用するためには、毎回「library」でパッケージを呼び出す必要があります。

パッケージのインストール方法：`install.packages(“パッケージ名”)`

パッケージを呼び出す方法：`library(パッケージ名)`



## STEP4 パッケージをインストールしよう



File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

R 4.3.1

Environment History Connections Tutorial

Import Dataset 178 MiB

Global Environment

Values

inu	10
neko	20

```
20 # -----
21 # Chapter02を進めるにあたり、はじめての場合は以下のパッケージをインストールしてください。
22 # パッケージのインストール
23 pkg <- c("tidyverse", "ggplot2", "openxlsx", "patchwork", "cmprsk", "survminer", "ggsurvfit", "finalfit", "survival")
24 install.packages(pkg)
25
26
27 # パッケージの呼び出し
28 library(tidyverse)
29 library(ggplot2)
30 library(openxlsx)
31 library(patchwork)
32 library(cmprsk)
33 library(survminer)
34 library(ggsurvfit)
35 library(finalfit)
36 library(survival)
37
38 # -----
39 # Chapter02 疫学指標
40
41 12:10 作成日 2023年6月3日
```

Console Terminal Background Jobs

R 4.3.1 · /cloud/project/

```
> c("RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」+「enter」")
[1] "RStudioでコードを実行したい場合は、コードが書かれた行にカーソルを合わせて「ctrl」+「enter」"
> c("複数行のコードを実行したければ、該当の行を選択して「ctrl」+「enter」")
[1] "複数行のコードを実行したければ、該当の行を選択して「ctrl」+「enter」"
> 3+7
[1] 10
> 2*2
[1] 4
> 2~10
2 ~ 10
> inu <- 10
> neko <- 20
> inu+neko
[1] 30
> # c("半角「#」のマークは、コードではなく、コメントとして認識される")
>
```

### パッケージをインストールして呼び出そう。

- 通常、「install.packages(“パッケージ名”)」を実行して、パッケージを呼び出しますが、23行目～24行目では、Chapter02で使用する以下のパッケージを一括でインストールしています。慣れなければ「install.packages(“パッケージ名”)」で構いません。
  - tidyverse
  - ggplot2
  - openxlsx
  - patchwork
  - cmprsk
  - survminer
  - ggsurvfit
  - survival
- 23行目、24行目のコードを実行します。
- コードを実行すると、コンソールにインストール過程が出力されますが、しばらく時間がかかるので、左下に「>」が表示されるまで何もせず待ちます。  
※30秒ほどで完了しますが、PCのスペックによっては、5分以上かかる可能性があります。

## STEP4 パッケージをインストールしよう

The screenshot shows the RStudio IDE with a script file named 'File2-3.R'. The script contains comments in Japanese and R code for installing and loading several packages. Lines 23 and 24 are highlighted in yellow, showing the `install.packages()` command. The console at the bottom shows the output of the installation process, with each package installation marked as 'DONE'. A yellow box highlights the console output, and a yellow arrow points from the highlighted code in the script to the console output. Another yellow arrow points from the console output to the second step of the instructions.

```
20 # -----
21 # Chapter02を進めるにあたり、はじめての場合は以下のパッケージをインストールしてください。
22 # パッケージのインストール
23 pkg <- c("tidyverse", "ggplot2", "openxlsx", "patchwork", "cmprsk", "survminer", "ggsurvfit", "finalfit", "survival")
24 install.packages(pkg)
25
26 # -----
27 # パッケージの呼び出し
28 library(tidyverse)
29 library(ggplot2)
30 library(openxlsx)
31 library(patchwork)
32 library(cmprsk)
33 library(survminer)
34 library(ggsurvfit)
35 library(finalfit)
36 library(survival)
37
38 # -----
39 # Chapter02 疫学指標
40
41 # (Untitled)
```

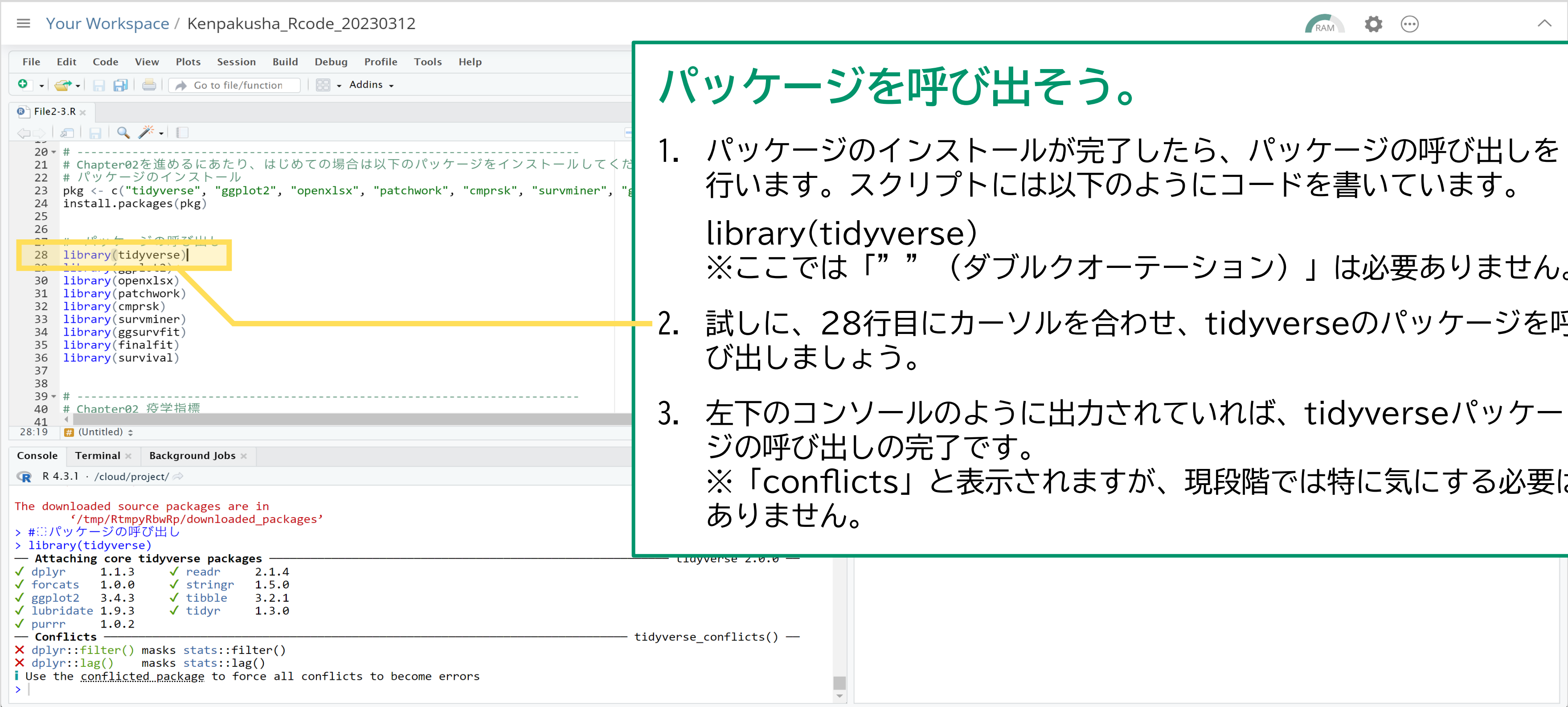
Environment: Global Environment  
Values:  
inu 10  
neko 20  
pkg chr [1:9] "tidyverse" "ggplot2" "openxlsx" "patchwork" "cmprsk" "surv..."

Console:  
R 4.3.1 > /cloud/project/  
\* installing \*binary\* package 'mitml' ...  
\* DONE (mitml)  
\* installing \*binary\* package 'rstatix' ...  
\* DONE (rstatix)  
\* installing \*binary\* package 'mice' ...  
\* DONE (mice)  
\* installing \*binary\* package 'ggpubr' ...  
\* DONE (ggpubr)  
\* installing \*binary\* package 'finalfit' ...  
\* DONE (finalfit)  
\* installing \*binary\* package 'survminer' ...  
\* DONE (survminer)  
The downloaded source packages are in  
'/tmp/RtmpyRbwRp/downloaded\_packages'  
>

### パッケージをインストールして呼び出そう。

- 通常、「install.packages(“パッケージ名”)」を実行して、パッケージを呼び出しますが、23行目～24行目では、Chapter02で使用する以下のパッケージを一括でインストールしています。慣れなければ「install.packages(“パッケージ名”)」で構いません。
  - tidyverse
  - ggplot2
  - openxlsx
  - patchwork
  - cmprsk
  - survminer
  - ggsurvfit
  - survival
- 23行目、24行目のコードを実行します。
- コードを実行すると、コンソールにインストール過程が出力されますが、しばらく時間がかかるので、左下に「>」が表示されるまで何もせず待ちます。  
※30秒ほどで完了しますが、PCのスペックによっては、5分以上かかる可能性があります。
- インストールが正常に終了すると、左下のコンソール画面のような状態になります。

## STEP4 パッケージをインストールしよう



The screenshot shows the RStudio IDE with a script file 'File2-3.R' open. The script contains R code for installing and loading several packages. A yellow box highlights the line `library(tidyverse)` at line 28. A yellow arrow points from this box to the second step of the instructions. The console at the bottom shows the output of the installation and loading process, including a list of core tidyverse packages and a warning about conflicts.

```
20 # -----
21 # Chapter02を進めるにあたり、はじめての場合は以下のパッケージをインストールしてくだ
22 # パッケージのインストール
23 pkg <- c("tidyverse", "ggplot2", "openxlsx", "patchwork", "cmprsk", "survminer", "g
24 install.packages(pkg)
25
26
27 # パッケージの呼び出し
28 library(tidyverse)
29 library(ggplot2)
30 library(openxlsx)
31 library(patchwork)
32 library(cmprsk)
33 library(survminer)
34 library(ggsurvfit)
35 library(finalfit)
36 library(survival)
37
38
39 # -----
40 # Chapter02 疫学指標
41
```

Console output:

```
R 4.3.1 · /cloud/project/

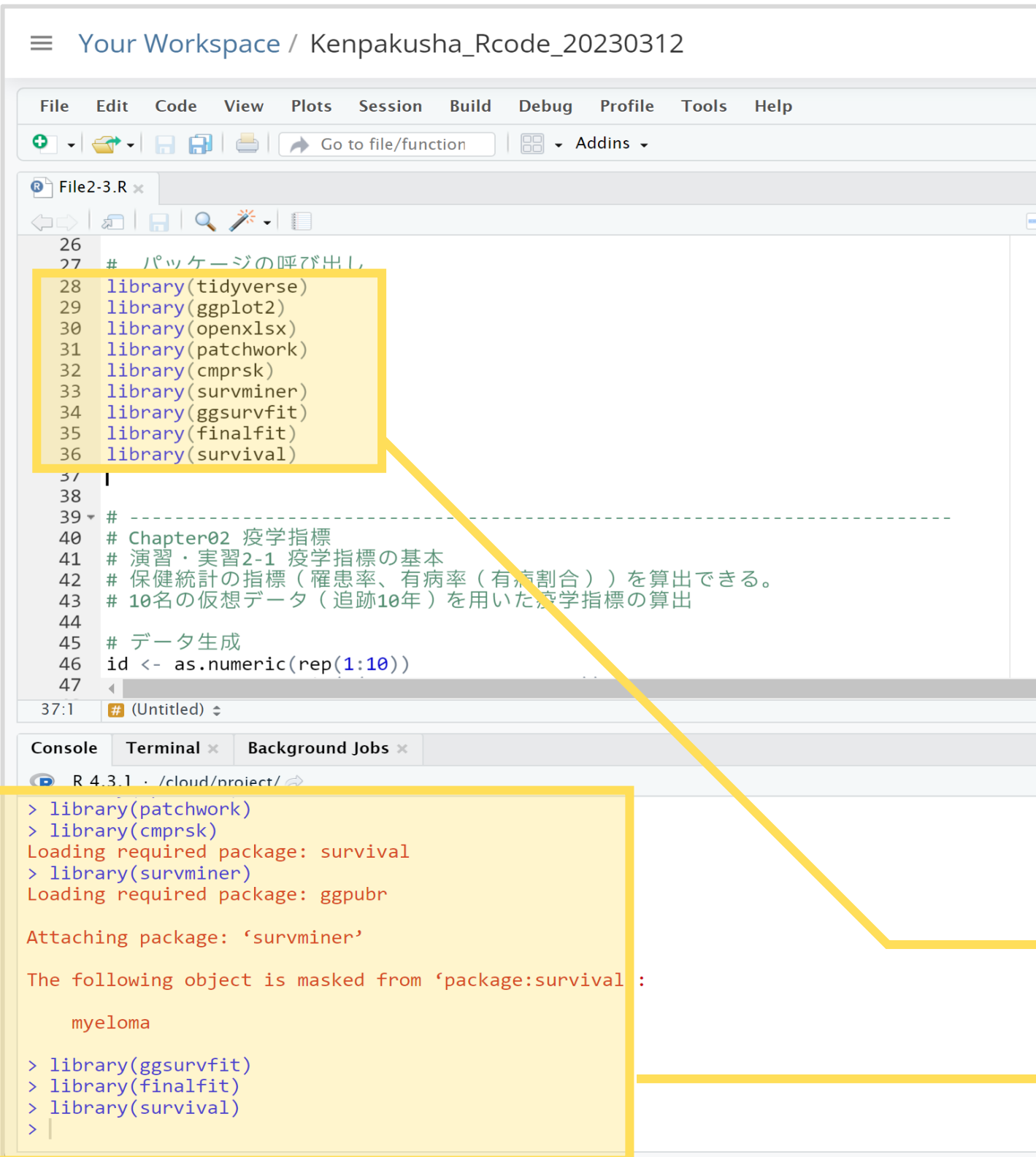
The downloaded source packages are in
  '/tmp/RtmpyRbwRp/downloaded_packages'
> # パッケージの呼び出し
> library(tidyverse)
Attaching core tidyverse packages
✓ dplyr      1.1.3    ✓ readr      2.1.4
✓ forcats   1.0.0    ✓ stringr   1.5.0
✓ ggplot2    3.4.3    ✓ tibble    3.2.1
✓ lubridate 1.9.3    ✓ tidyr     1.3.0
✓ purrr     1.0.2
--- Conflicts ---
tidyverse_conflicts()
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()     masks stats::lag()
i Use the conflicted...package to force all conflicts to become errors
>
```

### パッケージを呼び出そう。

1. パッケージのインストールが完了したら、パッケージの呼び出しを行います。スクリプトには以下のようにコードを書いています。  
`library(tidyverse)`  
※ここでは「” ”（ダブルクォーテーション）」は必要ありません。
2. 試しに、28行目にカーソルを合わせ、tidyverseのパッケージを呼び出しましょう。
3. 左下のコンソールのように出力されていれば、tidyverseパッケージの呼び出しの完了です。  
※「conflicts」と表示されますが、現段階では特に気にする必要はありません。



## STEP4 パッケージをインストールしよう



パッケージを呼び出そう。

1. パッケージのインストールが完了したら、パッケージの呼び出しを行います。スクリプトには以下のようにコードを書いています。  
`library(tidyverse)`  
※ここでは「” ”（ダブルクォーテーション）」は必要ありません。
2. 試しに、28行目にカーソルを合わせ、tidyverseのパッケージを呼び出しましょう。
3. 左下のコンソールのように出力されていれば、tidyverseパッケージの呼び出しの完了です。  
※「conflicts」と表示されますが、現段階では特に気にする必要はありません。
4. 同じように28行目から36行目までのコードも実行して、パッケージを呼び出しましょう。
5. ここまで完了したら、Chapter02を行うための準備完了です。

## STEP5 疫学指標を算出しよう

- 実習書の演習・実習2-1、2-2のワークシートで行った内容を、全てRStudio cloudで実施します。
- コードは既に記述済みなので、新たに追加したり編集したりする必要は一切ありません。順番に実行して、出力結果を確かめてください。
- 各コードで何の処理を行っているかは、スクリプトの中で「#」マークでコメントアウトしていますので、内容をよく確認してください。

# STEP5 疫学指標を算出しよう

≡ Your Workspace / Kenpakusha\_Rcode\_20230312

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins R 4.3.1

File2-3.R

Run Source

Environment History Connections Tutorial

Import Dataset 415 MiB

Global Environment

Values

inu	10
neko	20
pkg	chr [1:9] "tidyverse" "ggplot2" "openxlsx" "patchwork" "cmprsk" "surv...

```
40 # Chapter02 疫学指標
41 # 演習・実習2-1 疫学指標の基本
42 # 保健統計の指標（罹患率、有病率（有病割合））を算出できる。
43 # 10名の仮想データ（追跡10年）を用いた疫学指標の算出
44
45 # データ生成
46 id <- as.numeric(rep(1:10))
47 status <- as.numeric(c(1,0,0,0,0,1,0,1,1,0))
48 start_time <- as.numeric(c(0,0,1,5,2,2,0,0,1,0))
49 end_time <- as.numeric(c(7,10,10,10,10,5,10,8,4,10))
50 incidence_time <- as.numeric(c(3,1,0,0,9,4,0,3,2,2))
51 recovery_time <- as.numeric(c(0,2,0,0,10,0,0,0,0,5))
52 censor <- as.numeric(c(0,9,0,0,0,0,0,0,0,8))
53
54 sample_df <- data.frame(id, status, start_time, end_time,
55                         incidence_time, recovery_time, start_time, censor)
56
57 sample_df$status <- factor(sample_df$status,
58                           levels = c(0,1),
59                           labels = c("生存","死亡"))
60
61
```

Console Terminal Background Jobs

R 4.3.1 · /cloud/project/

> library(patchwork)
> library(cmprsk)
Loading required package: survival
> library(survminer)
Loading required package: ggpubr

Attaching package: 'survminer'

The following object is masked from 'package:survival':

 myeloma

> library(ggsurvfit)
> library(finalfit)
> library(survival)
>

## コホートの仮想データを生成します。

- id, status, start\_time, end\_time, incidence\_time, recovery\_time, censorのデータを生成し、sample\_dfという名前のデータフレームを作ります。
- sample\_dfにあるstatusという変数をfactor型の変数として定義しています。  
※変数の型は、いくつかの種類があります。代表的なものをいくつか挙げておきます。

numeric/integer :

実数値のデータ

logical :

論理型。TRUE/FALSEのどちらかの値。

character :

文字データ。

factor :

カテゴリカルなデータ。例えば男性を1, 女性を2

# STEP5 疫学指標を算出しよう

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

R 4.3.1

File2-3.R

```
46 id <- as.numeric(rep(1:10))
47 status <- as.numeric(c(1,0,0,0,0,1,0,1,1,0))
48 start_time <- as.numeric(c(0,0,1,5,2,2,0,0,1,0))
49 end_time <- as.numeric(c(7,10,10,10,10,5,10,8,4,10))
50 incidence_time <- as.numeric(c(3,1,0,0,9,4,0,3,2,2))
51 recovery_time <- as.numeric(c(0,2,0,0,10,0,0,0,0,5))
52 censor <- as.numeric(c(0,9,0,0,0,0,0,0,0,8))
53
54 sample_df <- data.frame(id, status, start_time, end_time,
55                         incidence_time, recovery_time, start_time, censor)
56 |
57 sample_df$status <- factor(sample_df$status,
58                             levels = c(0,1),
59                             labels = c("生存","死亡"))
60
61 glimpse(sample_df)
62 write_csv(sample_df, "data_sample.csv")
63
64 # 追跡状況の可視化
65 fig_cohort <- ggplot(sample_df) +
66   # 追跡期間の棒グラフ
67   # (Untitled)
```

Run

Source

Environment History Connections Tutorial

415 MiB

Global Environment

end_time	num [1:10]	7 10 10 10 10 5 10 8 4 10
id	num [1:10]	1 2 3 4 5 6 7 8 9 10
incidence_time	num [1:10]	3 1 0 0 9 4 0 3 2 2
inu		10

Console Terminal Background Jobs

R 4.3.1 · /cloud/project/

```
> sample_df$status <- factor(sample_df$status,
+                             levels = c(0,1),
+                             labels = c("生存","死亡"))
> glimpse(sample_df)
# A tibble: 10 x 8
#   id status start_time end_time incidence_time recovery_time start_time.1 censor
#   <dbl> <fct>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
1     1 死亡     0     7     3     0     0     0
2     2 生存     0     10    1     2     1     9
3     3 生存     0     10    0     0     0     0     0
4     4 生存     1     10    0     0     0     0     0
5     5 生存     5     10    9     4     0     0     0
6     6 死亡     2     5     0     0     0     0     0
7     7 死亡     2     10    0     0     0     0     0
8     8 死亡     0     8     3     0     0     0     0
9     9 死亡     0     4     2     0     0     0     0
10    10 生存     1     10    2     5     1     0     8
```

コホートの仮想データを生成します。

1. id, status, start\_time, end\_time, incidence\_time, recovery\_time, censor  
のデータを生成し、sample\_dfという名前のデータフレームを作ります。

2. sample\_dfにあるstatusという変数をfactor型の変数として定義しています。  
※変数の型は、いくつかの種類があります。代表的なものをいくつか挙げておきます。

numeric/integer：実数値のデータ

logical：論理型。TRUE/FALSEのどちらかの値。

character：文字データ。

factor：カテゴリカルなデータ。例えば男性を1, 女性を2

3. 61行目のglimpse(sample\_df)を実行すると、sample\_dfのデータフレームに含  
まれる内容を左下のコンソールで確認することができます。現在は10行×8列のデー  
タフレームであることがわかります。



## STEP5 疫学指標を算出しよう

### サンプルデータのファイルを保存します。

1. 「sample\_df」のデータフレームを出力したい（オフライン環境に保存したい）場合は、62行目の「write.csv」のコードを実行します。
2. 右下のペインに、「data\_sample.csv」のファイルを保存することができます。（オフライン環境に保存したファイルであれば、Excelなどで開くことができます）

```
48 start_time <- as.numeric(c(0,0,1,5,2,2,0,0,1,0))
49 end_time <- as.numeric(c(7,10,10,10,10,5,10,8,4,10))
50 incidence_time <- as.numeric(c(3,1,0,0,9,4,0,3,2,2))
51 recovery_time <- as.numeric(c(0,2,0,0,10,0,0,0,0,5))
52 censor <- as.numeric(c(0,9,0,0,0,0,0,0,0,8))
53
54 sample_df <- data.frame(id, status, start_time, end_time,
55                         incidence_time, recovery_time, start_time, censor)
56
57 sample_df$status <- factor(sample_df$status,
58                           levels = c(0,1),
59                           labels = c("生存","死亡"))
60
61 glimpse(sample_df)
62 write.csv(sample_df, "data_sample.csv")
63
64
65 # 追跡状況の可視化
66 fig_cohort <- ggplot(sample_df) +
67   #追跡期間の線を引く
68   annotate("segment", x = 4.5, y = 1, xend = 4.5, yend = 10.5,
```

①

```
Console Terminal Background Jobs
R 4.3.1 · /cloud/project/
+                               levels = c(0,1),
+                               labels = c("生存","死亡"))
>
> glimpse(sample_df)
Rows: 10
Columns: 8
$ id      <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
$ status  <fct> 死亡, 生存, 生存, 生存, 生存, 死亡, 生存, 死亡, 死亡, 生存
$ start_time <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0
$ end_time <dbl> 7, 10, 10, 10, 10, 5, 10, 8, 4, 10
$ incidence_time <dbl> 3, 1, 0, 0, 9, 4, 0, 3, 2, 2
$ recovery_time <dbl> 0, 2, 0, 0, 10, 0, 0, 0, 0, 5
$ start_time.1 <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0
$ censor     <dbl> 0, 9, 0, 0, 0, 0, 0, 0, 0, 8
> write.csv(sample_df, "data_sample.csv")
>
```

②

Files			Plots	Packages	Help	viewer	Presentation
New Folder New Blank File Upload Delete Rename More							
Cloud > project							
Name		Size	Modified				
..							
.Rhistory		0 B	Oct 9, 2023, 10:22 AM				
File2-3.R		13.8 KB	Oct 9, 2023, 2:21 PM				
project.Rproj		205 B	Oct 9, 2023, 2:38 PM				
File2-1.xlsx		339.4 KB	Oct 9, 2023, 2:42 PM				
File2-4.xlsx		1.4 MB	Oct 9, 2023, 2:42 PM				
File2-5.xlsx		62.3 KB	Oct 9, 2023, 2:42 PM				
File2-7.xlsx		294 KB	Oct 9, 2023, 2:42 PM				
data_sample.csv		377 B	Oct 9, 2023, 2:58 PM				

## STEP5 疫学指標を算出しよう

≡ Your Workspace / Kenpakusha\_Rcode\_20230312

RAM



File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

File2-3.R

```
64
65 # 追跡状況の可視化
66 fig_cohort <- ggplot(sample_df) +
67   #追跡期間の線を引く
68   annotate("segment", x = 4.5, y = 1, xend = 4.5, yend = 10.5,
69     linetype = 1, color = "grey", alpha = 0.7, arrow = arrow(length = unit(0.5, "cm")))+
70   geom_segment(data=subset(sample_df, incidence_time==0&ensor==0),
71     aes(x = start_time, xend = end_time, y = id, yend = id),
72     color = "grey", linewidth=1.8) +
73   geom_segment(data=subset(sample_df, incidence_time>=1&recovery_time==0),
74     aes(x = incidence_time, xend = end_time, y = id, yend = id),
75     color = "grey", linewidth=1, linetype="dashed") +
76   geom_segment(data=subset(sample_df, incidence_time>=1),
77     aes(x = start_time, xend = incidence_time, y = id, yend = id),
78     color = "grey", linewidth=1.8) +
79   geom_segment(data=subset(sample_df, incidence_time>=1&end_time==10&ensor==0),
80     aes(x = recovery_time, xend = end_time, y = id, yend = id),
81     color = "grey", linewidth=1.8) +
82   geom_segment(data=subset(sample_df, incidence_time>=1&end_time==10&ensor!=0),
83     aes(x = recovery_time, xend = censor, y = id, yend = id),
84     color = "grey", linewidth=1.8) +
85   geom_segment(data=subset(sample_df, incidence_time>=1&end_time==10),
86     aes(x = recovery_time, xend = end_time, y = id, yend = id),
87     color = "grey", linewidth=1.8)
```

Environment History Connections Tutorial

Import Dataset 417 MiB

R Global Environment

end_time	num [1:10]	7 10 10 10 10 5 10 8 4 10
id	num [1:10]	1 2 3 4 5 6 7 8 9 10
incidence_time	num [1:10]	3 1 0 0 9 4 0 3 2 2
inu		10

77:41 # (Untitled)

Console Terminal Background Jobs

R 4.3.1 /cloud/project/

```
+ levels = c(0,1),
+ labels = c("生存","死亡"))
>
> glimpse(sample_df)
Rows: 10
Columns: 8
$ id          <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
$ status      <fct> 死亡, 生存, 生存, 生存, 生存, 死亡, 生存, 死亡, 死亡, 生存
$ start_time  <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0
$ end_time    <dbl> 7, 10, 10, 10, 10, 5, 10, 8, 4, 10
$ incidence_time <dbl> 3, 1, 0, 0, 9, 4, 0, 3, 2, 2
$ recovery_time <dbl> 0, 2, 0, 0, 10, 0, 0, 0, 0, 5
$ start_time.1 <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0
$ censor      <dbl> 0, 9, 0, 0, 0, 0, 0, 0, 0, 8
> write.csv(sample_df, "data_sample.csv")
>
```

## 図を作成します。

1. せっかくなので、sample\_dfに含まれる仮想データを可視化したいと思います。実習書の図表2-1にあたります。
2. Rではデータの可視化のために、「ggplot2」という便利なパッケージがあります。
3. 66行目から126行目までがggplotのコードで、きれいにするために細かな調整を加えているので少しコードが長いですが、66行目から126行目のいずれかの場所にカーソルを合わせて、コードを実行してください。  
※該当部分のコードのすべてを選択する必要はありません。一つの塊のコードの、いずれかの行にカーソルがあれば該当のコードは実行できます。
4. ここで書いているコードの内容としては、ggplotによって作成した図を、「fig\_cohort」というオブジェクト名（箱というイメージ）で定義する（箱に入れる）、ということです。

## STEP5 疫学指標を算出しよう

### 図を作成します。

1. ここで書いているコードの内容としては、ggplotによって作成した図を、「fig\_cohort」というオブジェクト名（箱というイメージ）で定義する（箱に入れる）、ということです。
2. 129行目のコードを実行すると、右下のPlotsのペインに図が表示されます。  
※使用したデータをもとに、「Microsoft Excel」で同様の図を作るには限界がありますが、ggplotではデータの可視化においては、困ることがほぼありません。

≡ Your Workspace / Kenpakusha\_Rcode\_20230312

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

File2-3.R x

```
116 axis.title.y = element_text(face=bold, size = 25),
117 axis.text.x = element_text(size = 20),
118 axis.text.y = element_text(size = 20),
119 axis.ticks.length.x = unit(4, "pt"),
120 axis.ticks.length.y = unit(4, "pt"),
121 legend.text = element_text(size = 20),
122 plot.margin = unit(c(2,2,2,2), "lines"))+
123 labs(title = "",
124       x = "フォローアップ期間 (年)",
125       y = "ケース番号",
126       color="Status")
127
```

```
128 # 右下のPlotペインに図を表示
129 fig_cohort
```

```
130
131
132 # 追跡状況に関する図の保存 (Figure1) ②
```

```
133 ggsave(filename = "Fig1.jpeg",
134         plot = fig_cohort,
135         device = "jpeg",
136         width = 10,
137         height = 10)
138
```

132:1 # (Untitled) x

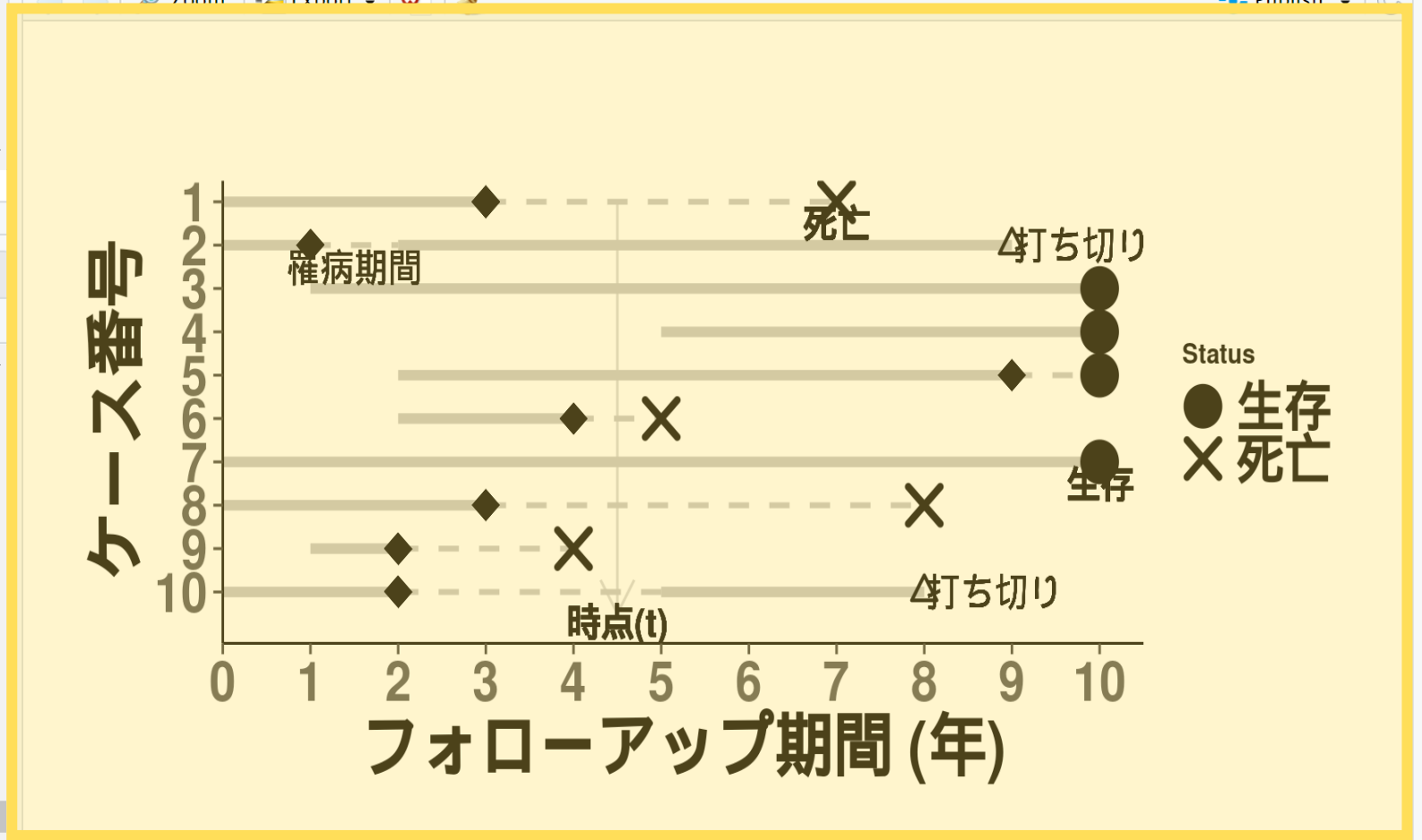
Console Terminal x Background Jobs x

R 4.3.1 · /cloud/project/

```
+ legend.position="right",
+ axis.title.x = element_text(face="bold", size = 25),
+ axis.title.y = element_text(face="bold", size = 25),
+ axis.text.x = element_text(size = 20),
+ axis.text.y = element_text(size = 20),
+ axis.ticks.length.x = unit(4, "pt"),
+ axis.ticks.length.y = unit(4, "pt"),
+ legend.text = element_text(size = 20),
+ plot.margin = unit(c(2,2,2,2), "lines"))+
+ labs(title = "",
+       x = "フォローアップ期間 (年)",
+       y = "ケース番号",
+       color="Status")
> # 右下のPlotペインに図を表示
> fig_cohort
>
```

fig_cohort	List of 9
sample_df	10 obs. of 8 variables
Values	
censor	num [1:10] 0 9 0 0 0 0 0 0 0 8
end_time	num [1:10] 7 10 10 10 10 5 10 8 4 10
id	num [1:10] 1 2 3 4 5 6 7 8 9 10
incidence_time	num [1:10] 3 1 0 0 9 4 0 3 2 2
inu	10

Files Plots Packages Help Viewer Presentation



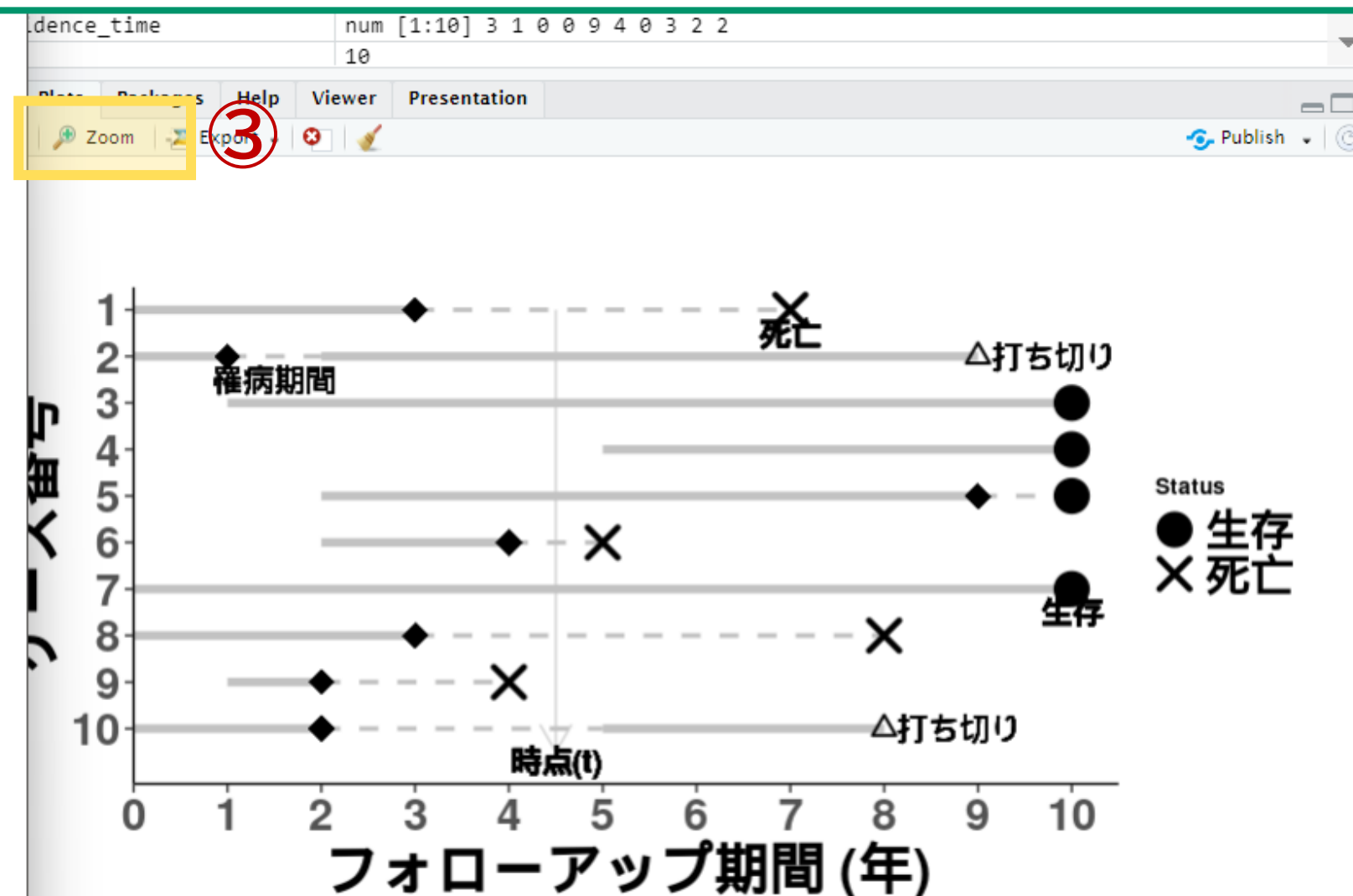
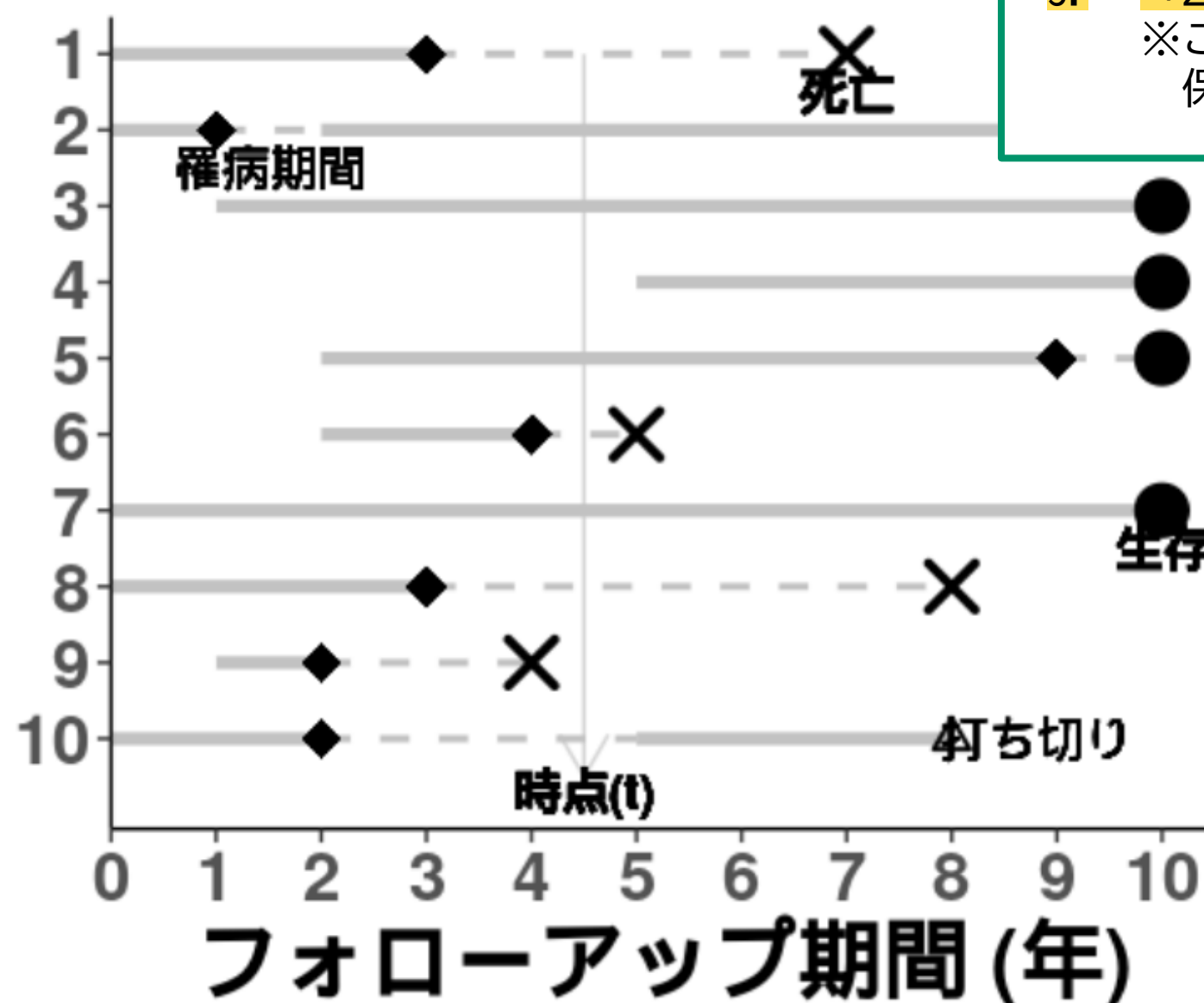


## STEP5 疫学指標を算出しよう

### 図を作成します。

1. ここで書いているコードの内容としては、ggplotによって作成した図を、「fig\_cohort」というオブジェクト名（箱というイメージ）で定義する（箱に入れる）、ということです。
2. 129行目のコードを実行すると、右下のPlotsのペインに図が表示されます。  
※使用したデータをもとに、「Microsoft Excel」で同様の図を作るには限界がありますが、ggplotではデータの可視化においては、困ることがほぼありません。
3. 「Zoom」をクリックすると、作成した図を拡大したウィンドが開きます。  
※この時点では、まだ図は“ファイル”として作成されていません。  
保存方法は後に説明します。

ケース番号





## STEP5 疫学指標を算出しよう

### ファイルを作成します。

1. 「fig\_cohort」をファイルとして保存します。133行目から139行目のggsaveのコードを実行すると、右下のFilesのペインに「Fig1.jpeg」が作成されます。

≡ Your Workspace / Kenpakusha\_Rcode\_20230312



File Edit Code View Plots Session Build Debug Profile Tools Help

+ - - - - - Go to file/function - - - - - Addins -

File2-3.R x

```
128 # 右下のPlotペインに図を表示
129 fig_cohort
130
131
132 # 追跡状況に関する図の保存 (Figure1)
133 ggsave(filename = "Fig1.jpeg",
134         plot = fig_cohort,
135         device = "jpeg",
136         width = 10,
137         height = 8,
138         unit = "in",
139         dpi = 600)
140
141
142 # 2) 罹患率の算出-----
143 # 人時法 (人年法) を用いた罹患率の算出
144 sample_df <- sample_df %>%
145   mutate(case = if_else(incidence_time >= 1, 1,0), # 疾患発生があれば1、そうでなければ0
146          pyear = if_else(incidence_time > 0,
147                          incidence_time-start_time,
148                          end time-start time)) %>% # 追跡期間を算出
149
140:1 # (Untitled) R Script
```

Console Terminal x Background Jobs x

```
R 4.3.1 · /cloud/project/
+   plot.margin = unit(c(2,2,2,2), "lines"))+
+   labs(title = "",
+        x = "フォローアップ期間 (年)",
+        y = "ケース番号",
+        color="Status")
> # 右下のPlotペインに図を表示
> fig_cohort
> # 追跡状況に関する図の保存 (Figure1)
> ggsave(filename = "Fig1.jpeg",
+         plot = fig_cohort,
+         device = "jpeg",
+         width = 10,
+         height = 8,
+         unit = "in",
+         dpi = 600)
>
```

censor	num [1:10]	0 9 0 0 0 0 0 0 0 8
end_time	num [1:10]	7 10 10 10 10 5 10 8 4 10
id	num [1:10]	1 2 3 4 5 6 7 8 9 10
incidence_time	num [1:10]	3 1 0 0 9 4 0 3 2 2
inu		10

Files Plots Packages Help Viewer Presentation

+ New Folder + New Blank File - Upload Delete Rename More -

Cloud > project

	Name	Size	Modified
	..		
	.Rhistory	0 B	Oct 9, 2023, 10:22 AM
	File2-3.R	13.8 KB	Oct 9, 2023, 2:21 PM
	project.Rproj	205 B	Oct 9, 2023, 2:38 PM
	File2-1.xlsx	339.4 KB	Oct 9, 2023, 2:42 PM
	File2-4.xlsx	1.4 MB	Oct 9, 2023, 2:42 PM
	File2-5.xlsx	62.3 KB	Oct 9, 2023, 2:42 PM
	File2-7.xlsx	294 KB	Oct 9, 2023, 2:42 PM
	data_sample.csv	377 B	Oct 9, 2023, 2:58 PM
	Fig1.jpeg	1.8 MB	Oct 9, 2023, 3:02 PM

## STEP5 疫学指標を算出しよう

### ファイルを作成します。

1. 「RStudio Cloud」上で作成し、「Files」のペインにあるファイルをPCのローカルストレージに保存するには、以下の順に作業します。
2. 右下のペインで、保存したいファイルのチェックボックスをクリックします。
3. 「More」をクリックします。
4. 「Export」をクリックします。

The screenshot displays the RStudio Cloud environment. The source pane on the left shows a file explorer with a list of files. The file 'Fig1.jpeg' is selected, indicated by a checked checkbox and a red circle with the number 2. The files pane on the right shows a list of files, including 'Fig1.jpeg', which is highlighted with a red circle with the number 2. A context menu is open over 'Fig1.jpeg', showing options like 'Copy...', 'Move...', and 'Export...'. The 'Export...' option is highlighted with a red circle with the number 4. The 'More' button in the files pane is also highlighted with a red circle with the number 3. The console at the bottom shows R code for plotting and saving the figure.

```
R 4.3.1 · /cloud/project/  
+ plot.margin = unit(c(2,2,2,2), "lines"))+  
+ labs(title = "",  
+       x = "フォローアップ期間 (年)",  
+       y = "ケース番号",  
+       color="Status")  
> # 右下のPlotペインに図を表示  
> fig_cohort  
> # 追跡状況に関する図の保存 (Figure1)  
> ggsave(filename = "Fig1.jpeg",  
+         plot = fig_cohort,  
+         device = "jpeg",  
+         width = 10,  
+         height = 8,  
+         unit = "in",  
+         dpi = 600)  
>
```

## STEP5 疫学指標を算出しよう

作成したファイルをPCにエクスポートします。

1. 中央に表示されるダイアログで、「Download」をクリックします。

The screenshot displays the RStudio interface. In the center, an 'Export Files' dialog box is open, asking the user to specify a name for the downloaded file. The name 'Fig1.jpeg' is entered in the text field. The 'Download' button is highlighted with a red circle and the number 1. The background shows the RStudio workspace with a script editor containing R code for saving a plot and calculating incidence rates, a console window, and a file explorer showing the project files.

```
# 追跡状況に関する図の保存 (Figure1)
ggsave(filename = "Fig1.jpeg",
        plot = fig_cohort,
        device = "jpeg",
        width = 10,
        height = 8,
        unit = "in",
        dpi = 600)

# 2) 罹患率の算出-----
# 人時法 (人年法) を用いた罹患率の算出
sample_df <- sample_df %>%
  mutate(case = if_else(incidence_time >= 1, 1, 0), # 疾患発生があれば1、そうでなければ0
         pyear = if_else(incidence_time > 0,
                        incidence_time-start_time,
                        end time-start time)) %>% # 追跡期間を算出
```

Name	Size	Modified
story	0 B	Oct 9, 2023, 10:22 AM
2-3.R	13.8 KB	Oct 9, 2023, 2:21 PM
project.Rproj	205 B	Oct 9, 2023, 2:38 PM
File2-1.xlsx	339.4 KB	Oct 9, 2023, 2:42 PM
File2-4.xlsx	1.4 MB	Oct 9, 2023, 2:42 PM
File2-5.xlsx	62.3 KB	Oct 9, 2023, 2:42 PM
File2-7.xlsx	294 KB	Oct 9, 2023, 2:42 PM
data_sample.csv	377 B	Oct 9, 2023, 2:58 PM
Fig1.jpeg	1.8 MB	Oct 9, 2023, 3:02 PM

## STEP5 疫学指標を算出しよう

作成したファイルをPCにエクスポートします。

1. 中央に表示されるダイアログで、「Download」をクリックします。
2. PCのダウンロードフォルダに保存されます。
3. クリックします（表示される場所は、お使いのPCの環境によって変わるかもしれません）。

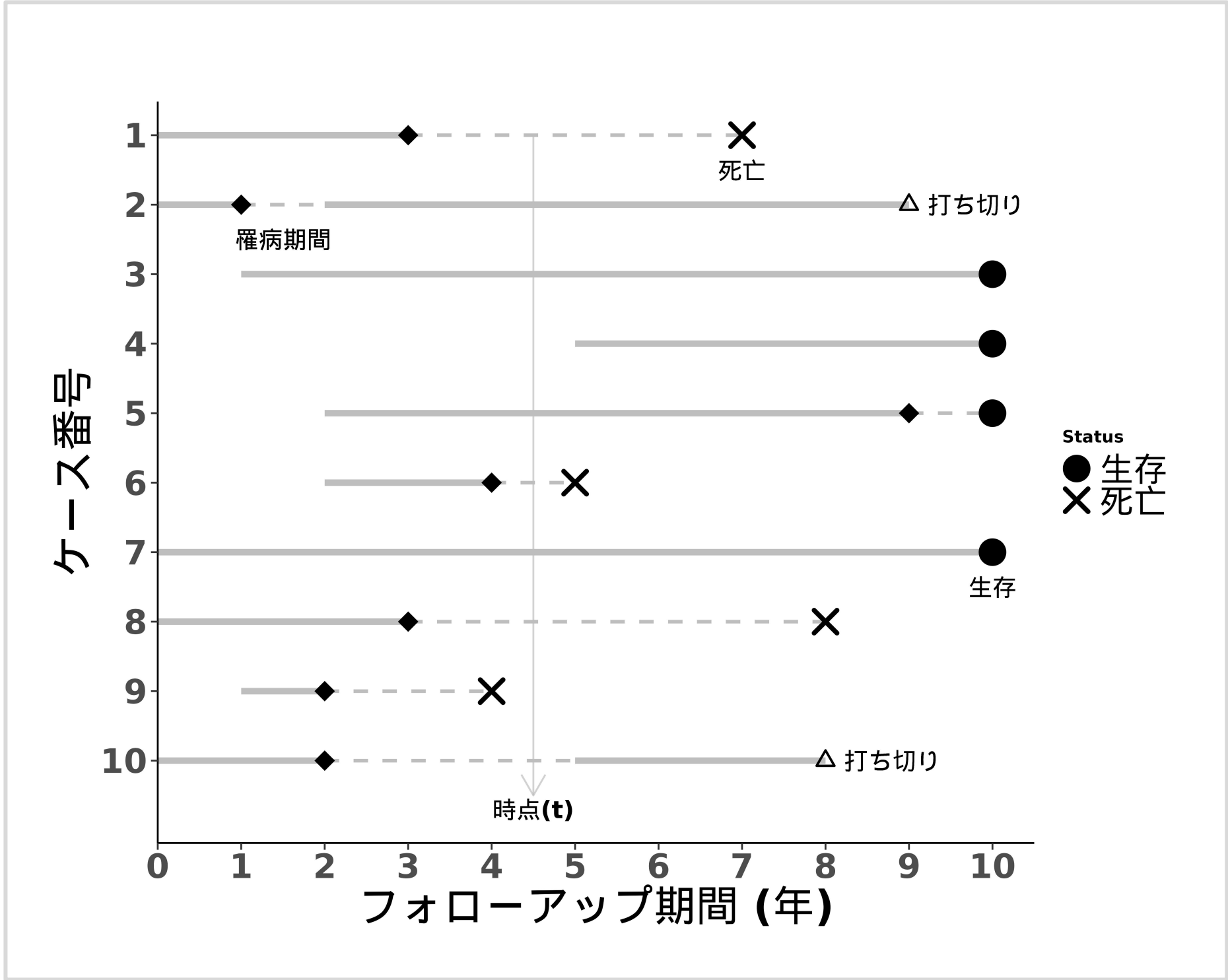
The screenshot displays the RStudio environment. The console window at the bottom shows the following R code being executed:

```
R 4.3.1 · /cloud/project/
+ plot.margin = unit(c(2,2,2,2), "lines"))+
+ labs(title = "",
+       x = "フォローアップ期間 (年)",
+       y = "ケース番号",
+       color="Status")
> # 右下のPlotペインに図を表示
> fig_cohort
> # 追跡状況に関する図の保存 (Figure1)
> ggsave(filename = "Fig1.jpeg",
+         plot = fig_cohort,
+         device = "jpeg",
+         width = 10,
+         height = 8,
+         unit = "in",
+         dpi = 600)
>
```

The Files pane on the right shows a list of files in the 'project' directory. The file 'Fig1.jpeg' is highlighted, and a yellow callout box with a red circle containing the number 3 points to the 'Download' button in the Files pane.

Name	Size	Modified
..		
.Rhistory	0 B	Oct 9, 2023, 10:22 AM
File2-3.R	13.8 KB	Oct 9, 2023, 2:21 PM
project.Rproj	205 B	Oct 9, 2023, 2:38 PM
File2-1.xlsx	339.4 KB	Oct 9, 2023, 2:42 PM
File2-4.xlsx	1.4 MB	Oct 9, 2023, 2:42 PM
File2-5.xlsx	62.3 KB	Oct 9, 2023, 2:42 PM
File2-7.xlsx	294 KB	Oct 9, 2023, 2:42 PM
data_sample.csv	377 B	Oct 9, 2023, 2:58 PM
Fig1.jpeg	1.8 MB	Oct 9, 2023, 3:02 PM

STEP5 疫学指標を算出しよう



作成した図のファイルを開きます。

1. 実習書の図表2-1を作成することができました。



## STEP5 疫学指標を算出しよう

RAM

⚙️ ⋮

⌵

File Edit Code View Plots Session Build Debug Profile Tools Help

+ ↩️ ↪️ 📄 📄 📄 📄 📄 Go to file/function Addins R 4.3.1

File2-3.R x

138 unit = "in",  
139 dpi = 600)  
140  
141  
142 # 2) 罹患率の算出-----  
143 # 人時法（人年法）を用いた罹患率の算出  
144 sample\_df <- sample\_df %>%  
145 mutate(case = if\_else(incidence\_time >= 1, 1, 0), # 疾患発生があれば1、そうでなければ0  
146 pyear = if\_else(incidence\_time > 0,  
147 incidence\_time-start\_time,  
148 end\_time-start\_time)) %>% # 追跡期間を算出  
149 arrange(id) # idで並び替え  
150  
151 glimpse(sample\_df) # データフレームの確認  
152 IR <- sample\_df %>% summarise(IR = sum(case)/sum(pyear)) # 罹患率0.163（時^-1）  
153 IR  
154  
155 # 3) 有病率（有病割合）の算出-----  
156 p.prop <- sample\_df %>%  
157 mutate(time\_point = 4.5,  
158 case.exist =  
159 case\_when(incidence\_time==0 ~ 0,  
160 recovery\_time==0 & time\_point<=end\_time & time\_point>=incidence\_time ~ 1,  
161 else ~ 0)  
162  
154:1 # 2) 罹患率の算出

Console Terminal x Background Jobs x

R 4.3.1 · /cloud/project/

Rows: 10  
Columns: 10  
\$ id <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10  
\$ status <fct> 死亡, 生存, 生存, 生存, 生存, 死亡, 生存, 死亡, 死亡, 生存  
\$ start\_time <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0  
\$ end\_time <dbl> 7, 10, 10, 10, 10, 5, 10, 8, 4, 10  
\$ incidence\_time <dbl> 3, 1, 0, 0, 9, 4, 0, 3, 2, 2  
\$ recovery\_time <dbl> 0, 2, 0, 0, 10, 0, 0, 0, 0, 5  
\$ start\_time.1 <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0  
\$ censor <dbl> 0, 9, 0, 0, 0, 0, 0, 0, 0, 8  
\$ case <dbl> 1, 1, 0, 0, 1, 1, 0, 1, 1, 1  
\$ pyear <dbl> 3, 1, 9, 5, 7, 2, 10, 3, 1, 2  
> IR <- sample\_df %>% summarise(IR = sum(case)/sum(pyear)) # 罹患率0.163（時^-1）  
> IR  
IR  
1 0.1627907  
>

### 罹患率を算出します。

- データ生成が終わったので、疫学指標の算出に入ります。
- 144行目から149行目では、「sample\_df」のデータを用いて罹患率を算出するために、caseとpyearの計算をしています。
- 152行目のコードを実行して罹患率を算出します。

	Name	Size	Modified
<input type="checkbox"/>	..		
<input type="checkbox"/>	.Rhistory	0 B	Oct 9, 2023, 10:22 AM
<input type="checkbox"/>	File2-3.R	13.8 KB	Oct 9, 2023, 2:21 PM
<input type="checkbox"/>	project.Rproj	205 B	Oct 9, 2023, 2:38 PM
<input type="checkbox"/>	File2-1.xlsx	339.4 KB	Oct 9, 2023, 2:42 PM
<input type="checkbox"/>	File2-4.xlsx	1.4 MB	Oct 9, 2023, 2:42 PM
<input type="checkbox"/>	File2-5.xlsx	62.3 KB	Oct 9, 2023, 2:42 PM
<input type="checkbox"/>	File2-7.xlsx	294 KB	Oct 9, 2023, 2:42 PM
<input type="checkbox"/>	data_sample.csv	377 B	Oct 9, 2023, 2:58 PM
<input checked="" type="checkbox"/>	Fig1.jpeg	1.8 MB	Oct 9, 2023, 3:02 PM

## STEP5 疫学指標を算出しよう

File2-3.R

```
138   unit = "m",
139   dpi = 600)
140
141
142 # 2) 罹患率の算出-----
143 # 人時法（人年法）を用いた罹患率の算出
144 sample_df <- sample_df %>%
145   mutate(case = if_else(incidence_time >= 1, 1, 0), # 疾患発生があれば1、そうでなければ0
146     pyear = if_else(incidence_time > 0,
147       incidence_time-start_time,
148       end_time-start_time)) %>% # 追跡期間を算出
149   arrange(id) # idで並び替え
150
151 glimpse(sample_df) # データフレームの確認
152 IR <- sample_df %>% summarise(IR = sum(case)/sum(pyear)) # 罹患率0.163（時^-1）
153 IR
154
155 # 3) 有病率（有病割合）の算出-----
156 p.prop <- sample_df %>%
157   mutate(time_point = 4.5,
158     case.exist =
159       case_when(incidence_time==0 ~ 0,
160         recovery_time==0 & time_point<=end_time & time_point>=incidence_time ~ 1,
161         TRUE ~ 0))
154:1 # 2) 罹患率の算出
```

Environment

History

Connections

Tutorial

Console

Terminal

Background Jobs

R 4.3.1 · /cloud/project/

Rows: 10  
Columns: 10  
\$ id <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10  
\$ status <fct> 死亡, 生存, 生存, 生存, 生存, 死亡, 生存, 死亡, 死亡, 生存  
\$ start\_time <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0  
\$ end\_time <dbl> 7, 10, 10, 10, 10, 5, 10, 8, 4, 10  
\$ incidence\_time <dbl> 3, 1, 0, 0, 9, 4, 0, 3, 2, 2  
\$ recovery\_time <dbl> 0, 2, 0, 0, 10, 0, 0, 0, 0, 5  
\$ start\_time.1 <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0  
\$ censor <dbl> 0, 9, 0, 0, 0, 0, 0, 0, 0, 8  
\$ case <dbl> 1, 1, 0, 0, 1, 1, 0, 1, 1, 1  
\$ pyear <dbl> 3, 1, 9, 5, 7, 2, 10, 3, 1, 2  
> IR <- sample\_df %>% summarise(IR = sum(case)/sum(pyear)) # 罹患率0.163（時^-1）  
> IR  
#> # A tibble: 1 x 1  
#> IR  
1 0.1627907  
>

File2-1.xlsx

File2-4.xlsx

File2-5.xlsx

File2-7.xlsx

data\_sample.csv

Fig1.jpeg

339.4 KB

1.4 MB

62.3 KB

294 KB

377 B

1.8 MB

Oct 9, 2023, 2:42 PM

Oct 9, 2023, 2:42 PM

Oct 9, 2023, 2:42 PM

Oct 9, 2023, 2:42 PM

Oct 9, 2023, 2:58 PM

Oct 9, 2023, 3:02 PM

### 罹患率を算出します。

- データ生成が終わったので、疫学指標の算出に入ります。
- 144行目から149行目では、「sample\_df」のデータを用いて罹患率を算出するために、caseとpyearの計算をしています。
- 152行目のコードを実行して罹患率を算出します。
- 左下のコンソールに結果が出力されます。罹患率は約0.163となりました。

4

# STEP6 残りのRコードを1つずつ実行して 出力結果を確認しよう

- これ以降はRファイルのスク립トの記述に従って、同じように各コードを実行してください。実習書のChapter02で実施した内容をR言語を使って再現することができます。

The screenshot displays the RStudio interface with the following components:

- Code Editor:** Contains R code for calculating the incidence rate (IR) and prevalence (p.prop). The code includes data manipulation steps using `mutate`, `arrange`, and `summarise`.
- Environment:** Shows the current environment with variables `fig_cohort` (List of 9), `IR` (1 obs. of 1 variable), and `sample_df` (10 obs. of 10 variables).
- Files:** Lists the project files, including `File2-3.R`, `project.Rproj`, `File2-1.xlsx`, `File2-4.xlsx`, `File2-5.xlsx`, `File2-7.xlsx`, `data_sample.csv`, and `Fig1.jpeg`.
- Console:** Shows the output of the R code, including the structure of the `sample_df` data frame and the calculated IR value of 0.1627907.

```
# 2) 罹患率の算出-----
# 人時法（人年法）を用いた罹患率の算出
sample_df <- sample_df %>%
  mutate(case = if_else(incidence_time >= 1, 1, 0), # 疾患発生があれば1、そうでなければ0
         pyear = if_else(incidence_time > 0,
                        incidence_time - start_time,
                        end_time - start_time)) %>% # 追跡期間を算出
  arrange(id) # idで並び替え

glimpse(sample_df) # データフレームの確認
IR <- sample_df %>% summarise(IR = sum(case)/sum(pyear)) # 罹患率0.163（時^-1）
IR

# 3) 有病率（有病割合）の算出-----
p.prop <- sample_df %>%
  mutate(time_point = 4.5,
         case.exist =
           case_when(incidence_time==0 ~ 0,
                    recovery_time==0 & time_point<=end_time & time_point>=incidence_time ~ 1,
                    TRUE ~ 0))
p.prop
```

Console Output:

```
R 4.3.1 > /cloud/project/
Rows: 10
Columns: 10
 $ id      <dbl> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
 $ status  <fct> 死亡, 生存, 生存, 生存, 生存, 死亡, 生存, 死亡, 死亡, 生存
 $ start_time <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0
 $ end_time  <dbl> 7, 10, 10, 10, 10, 5, 10, 8, 4, 10
 $ incidence_time <dbl> 3, 1, 0, 0, 9, 4, 0, 3, 2, 2
 $ recovery_time <dbl> 0, 2, 0, 0, 10, 0, 0, 0, 0, 5
 $ start_time.1 <dbl> 0, 0, 1, 5, 2, 2, 0, 0, 1, 0
 $ censor      <dbl> 0, 9, 0, 0, 0, 0, 0, 0, 0, 8
 $ case        <dbl> 1, 1, 0, 0, 1, 1, 0, 1, 1, 1
 $ pyear       <dbl> 3, 1, 9, 5, 7, 2, 10, 3, 1, 2
> IR <- sample_df %>% summarise(IR = sum(case)/sum(pyear)) # 罹患率0.163（時^-1）
> IR
      IR
1 0.1627907
>
```

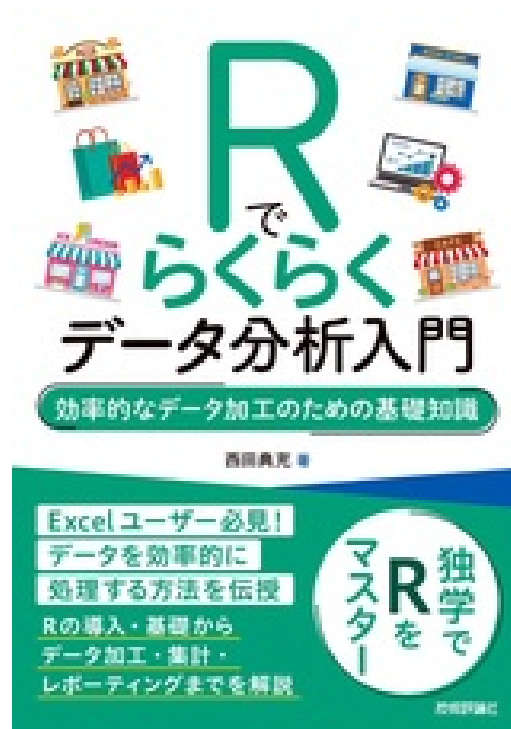
# STEP7 R言語を使って、自由にデータ処理・解析できるようになる(参考資料紹介)

- R言語に関する書籍は様々ありますが、必ず役立ついくつかの情報を紹介いたします。最初はコードを書くことに慣れないかもしれませんが、ここに挙げた書籍やHPを終える頃には、R言語の快適さ・便利さに気づくと思います。



・1冊目としておすすめ

金芳堂  
定価 3,520円 (税込み)



・もう一步レベルアップするために

技術評論社  
定価 3,278円 (税込み)

## ■ Rによるデータ処理の学習に役立つHP情報

### 1. 疫学のためのRハンドブック

<https://epirhandbook.com/jp/index.html>

(疫学者や疫学実務者によって書かれた、疫学者や疫学実務者のためのハンドブックです。基本編から実践的な内容のすべてが記載されています)

### 2. 私たちのR

<https://www.jaysong.net/RBook/index.html>

(宋財沄 (SONG Jaehyun) 先生と 矢内勇生先生が共同で執筆しているHPです。R言語の扱い方について細かく記載されています。)





## まとめ

- RStudioはR言語を扱いやすくした統合開発環境であり、疫学分野で汎用されています。そのクラウド環境がRStudio cloudです。
- あらゆるデータ処理、可視化、統計解析等を実行できます。
- 処理した内容は全てRファイルにコードとして記録が残るため、データ処理の再現性を保つことができます。
- 大量のデータ解析も、わずかな時間で処理できます。
- とにかく便利です。