

自分で触ってよくわかる

# 拡張パッケージの話：

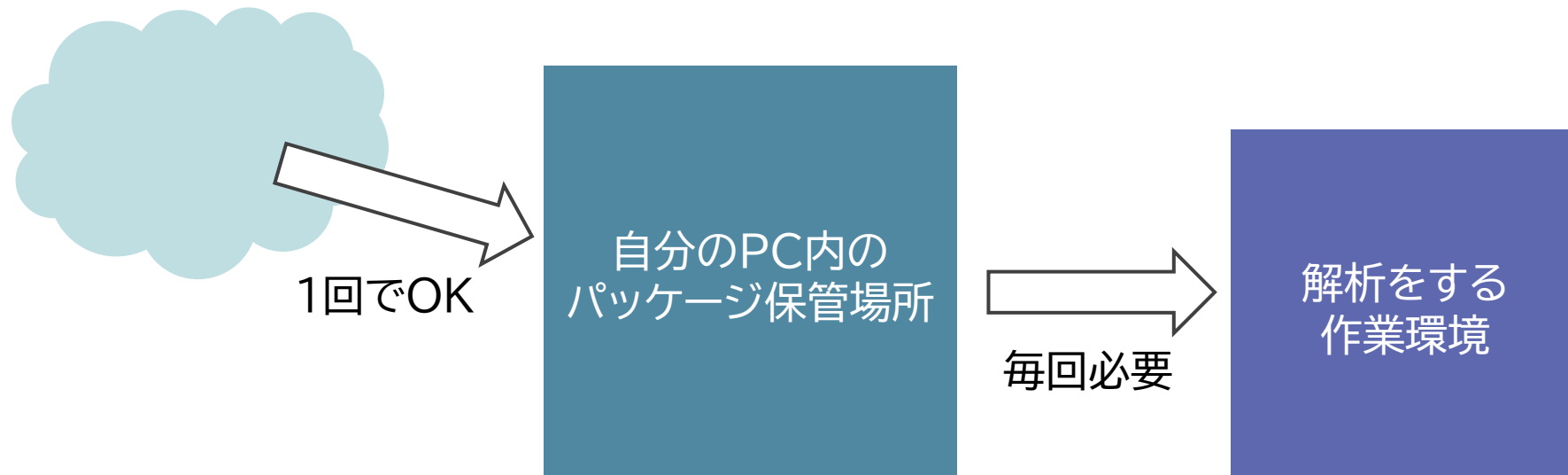
特化した機能を追加する

神戸市立医療センター中央市民病院  
臨床研究推進センター

宮越 千智

# 拡張パッケージとは

- 拡張パッケージをダウンロードすることで特定の機能を追加することができる  
(Pythonではライブラリと呼ぶ方が適切かもしれませんが、ここでは用語を厳密に使い分けていません)
- 2段階の手順が必要：
  1. パッケージのインストール: インターネット経由でダウンロード(1回行えばOK)
  2. 解析環境への読み込み: 解析を始めるときには毎回読み込む必要あり



Rを使いたい人 

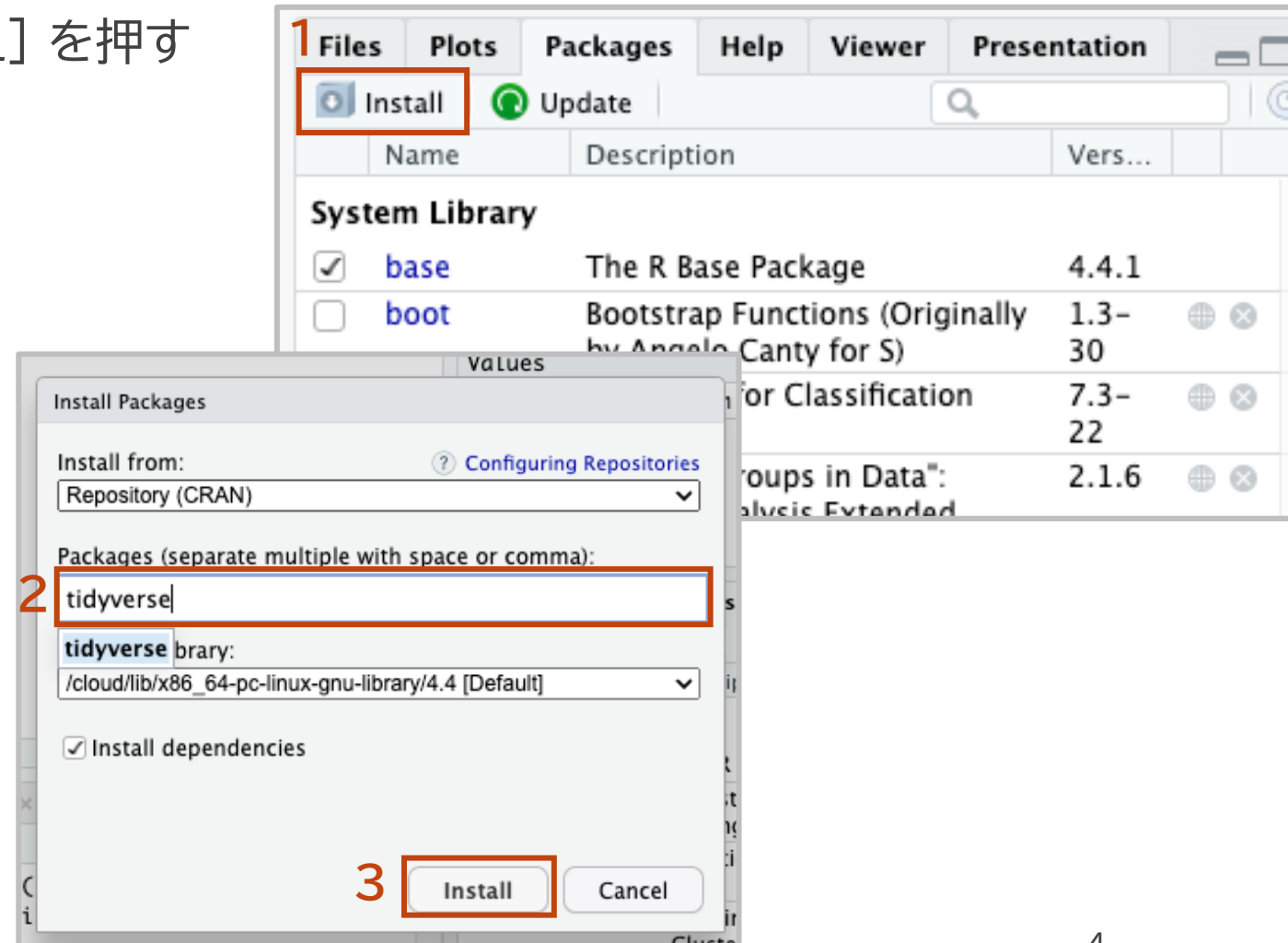
# Rを使いたい人: パッケージをインストールする(1回でOK)

1. 右下の [Packages] タブで [Install] を押す
2. パッケージ名を入力する  
(頭文字を入力すると候補が提案されます)
3. [Install] を押す

## 別法:

- 以下のコードを実行する

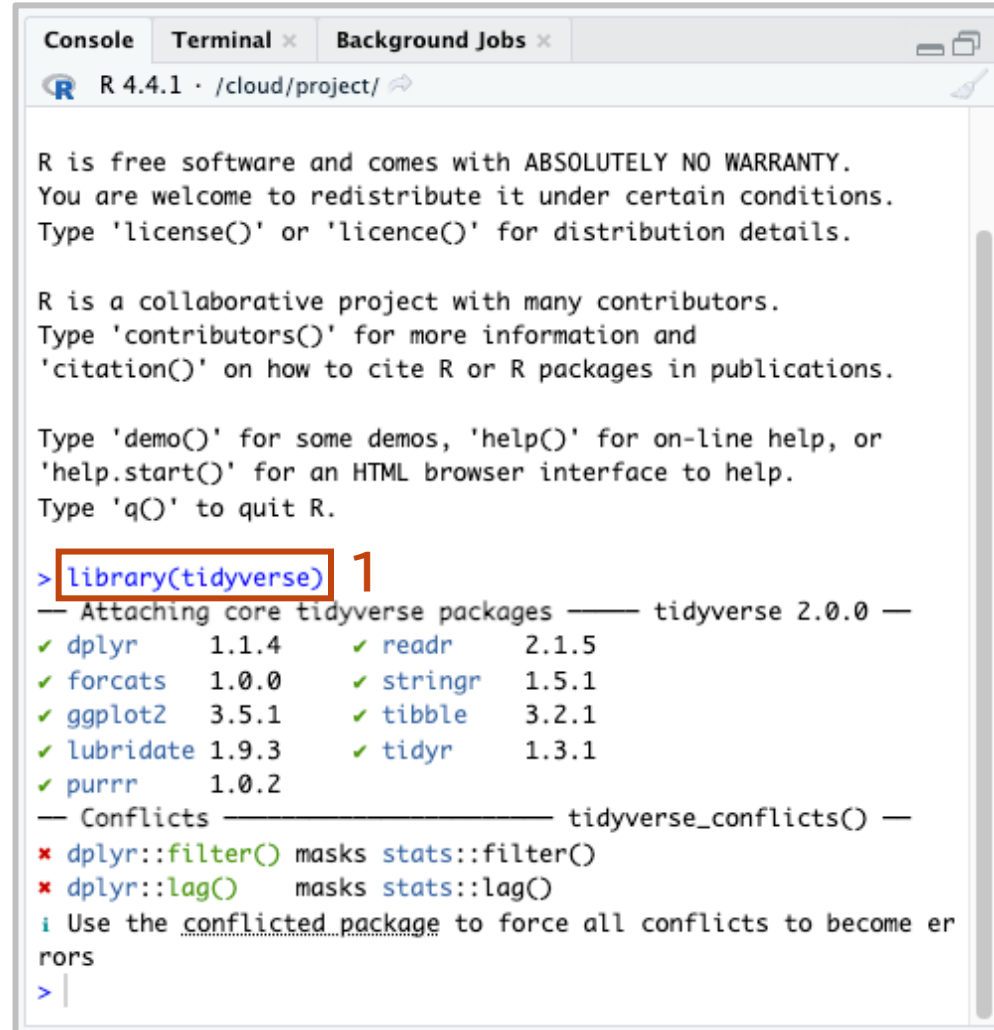
```
install.packages("パッケージ名")
```



# Rを使いたい人: パッケージを読み込む(毎回必要)

1. 以下のコードを実行する

```
library(パッケージ名)
```



```
Console Terminal Background Jobs
R 4.4.1 - /cloud/project/

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.

> library(tidyverse) 1
— Attaching core tidyverse packages — tidyverse 2.0.0 —
✓ dplyr      1.1.4      ✓ readr      2.1.5
✓ forcats    1.0.0      ✓ stringr    1.5.1
✓ ggplot2    3.5.1      ✓ tibble     3.2.1
✓ lubridate  1.9.3      ✓ tidyr      1.3.1
✓ 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
> |
```

# Rを使いたい人: 頻用パッケージ一覧

## tidyverse


- ✓ データハンドリングに役立つパッケージ(dplyr, tidyrなど)や強力な可視化パッケージ(ggplot2)をまとめたパッケージ

## gtsummary

- ✓ 記述要約や回帰分析の結果を簡単にまとめて表にしてくれるパッケージ

## survival

- ✓ 生存時間解析を行うときには必須のパッケージ

Pythonを使いたい人 



# Pythonを使いたい人: ライブラリをインストールして読み込む

1. 以下のコードを実行してインストールする(1回でOK)

```
!pip install パッケージ名
```

- ✓ 既にインストール済みのものもあります  
(例のnumpyはインストール済み)

```
+ コード + テキスト
```

```
[6] 1 !pip install numpy 1
```

```
{x} Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (1.25.2)
```

```
0 秒 1 import numpy as np 2
```

2. 以下のコードを実行して読み込む(毎回必要)

```
import パッケージ名 as 略名
```

- ✓ 略名を指定しておくと、毎回ライブラリのフルネームを使わなくていい





# Pythonを使いたい人: 頻用ライブラリー一覧

## **numpy**(略記: np)

- ✓ 数値計算の基本ライブラリ

## **pandas**(略記: pd)

- ✓ データフレームを用いてデータ管理するためのライブラリ

## **matplotlib**(略記: plt)

- ✓ データ可視化ライブラリ

## **scikit-learn**(略記: sklearn)

- ✓ 機械学習にデータ前処理、モデル構築、評価が一貫して行えるライブラリ

## 課題2：四則演算など

- 身長170cm、体重65kgの人のBMIを計算してみましょう  
( $\text{BMI} = \text{体重}[\text{kg}] \div (\text{身長}[\text{m}])^2$ です)
- 上記のBMIの自然対数( $=\log_e$ )を計算してみましょう

\* Rは標準でも上記の計算ができますが、Pythonではnumpy(など)を読み込む必要があります。

# よく使う計算記号

計算		例	R	Python (import numpy as np)	結果
足し算		$2+3$	$2+3$		5
引き算		$5-1$	$5-1$		4
掛け算		$2\times 9$	$2*9$		18
割り算	商	$11\div 3$	$11/3$		3.66...
	商の整数部		$11 \ \%/\% \ 3$	$11 \ /\!/\ 3$	3
	余り		$11 \ \% \% \ 3$	$11 \ \% \ 3$	2
べき乗		$2^5$	$2^5$ または $2^{**}5$	$2^{**}5$ または $\text{pow}(2,5)$	32
指数・対数	ネイピア数のべき数	$e^2$	$\text{exp}(2)$	$\text{np.exp}(2)$	7.389...
	自然対数	$\log_e 2$	$\log(2)$	$\text{np.log}(2)$	0.693...



# Pythonを使いたい人： なぜ`exp()`ではなく`np.exp()`？

- Pythonは関数が含まれているライブラリを明示する必要がある
  - ✓ Rでも同じ名前の関数が複数のパッケージに含まれていることがある
  - ✓ 使い分けたいときはコロンを2つ(`::`)を用いる  
例: `dplyr::filter()`, `stats::filter()`
- ピリオド(`.`)は「～に含まれる」「～に属する」という意味
  - ✓ 例: npライブラリに含まれる`exp()`関数
- クラスやオブジェクトの属性・メソッドにアクセスするときにも用いられる

# 今回のまとめ

- ✓ パッケージは全て無料です
- ✓ RやPythonでやりたいことがあったら、  
まずはどのパッケージを使ったらいいか調べましょう