Rッケージをつくる

<u>@</u>Tokyo.R #119, Sep 20, 2025



R パッケージを作りました!

- <u>日本アクチュアリー会</u>の研究活動の一環で作成した R パッケージ <u>midr</u> が CRAN に公開 されました。

midr: Learning from Black-Box Models by Maximum Interpretation Decomposition

The goal of 'midr' is to provide a model-agnostic method for interpreting and explaining black-box predictive models by creating a globally interpretable surrogate model. The package implements 'Maximum Interpretation Decomposition' (MID), a functional decomposition technique that finds an optimal additive approximation of the original model. This approximation is achieved by minimizing the squared error between the predictions of the black-box model and the surrogate model. The theoretical foundations of MID are described in Iwasawa & Matsumori (2025) [Forthcoming], and the package itself is detailed in Asashiba et al. (2025) <doi:10.48550/arXiv.2506.08338>.

Version: 0.5.2

Imports: graphics, grDevices, <u>Rcpp</u>, <u>RcppEigen</u>, <u>rlang</u>, stats, utils

LinkingTo: <u>Rcpp</u>, <u>RcppEigen</u>

Suggests: datasets, ggplot2, khroma, knitr, RColorBrewer, rmarkdown, scales, shapviz, testthat, viridisLite

Published: 2025-09-07

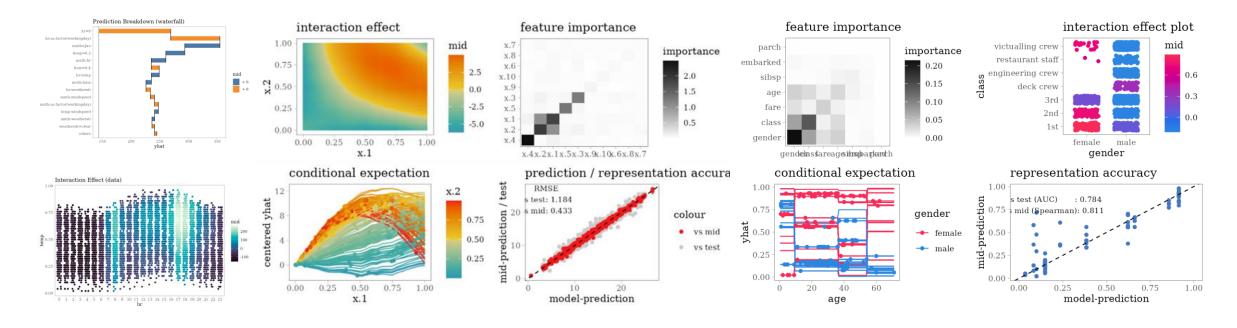
DOI: <u>10.32614/CRAN.package.midr</u>

Author: Ryoichi Asashiba [aut, cre], Hirokazu Iwasawa [aut], Reiji Kozuma [ctb]



本日共有させていただきたいこと

- midr は、機械学習モデルの解釈を目的とするパッケージです。
 - 「ブラックボックスモデルの予測値を予測する代理モデルを構築する機能」と「その代理モデルを様々な切り口で可視化する機能」を実装しています(PDP、ALE、SHAP、PFIなどと同様の用途)。



- パッケージを宣伝したい!...というのが、今日のセッションへの本当の応募動機です。 ...ですが、今日は R パッケージをつくった経験に焦点をあててお話させていただきます。

R パッケージをつくる

@Tokyo.R #119, Sep 20, 2025



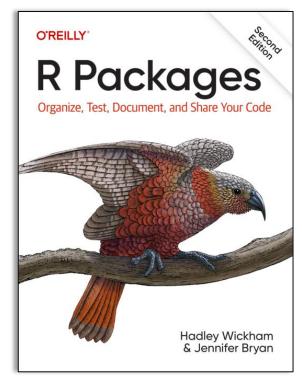
R Packages (2e) を読めば十分…?

- Rパッケージを自作するために必要な情報はすべて R Packages (2e) に書かれています

…ですが、この本の説明はとてもしっかりした作法にのっとっていて、完璧に従うのはけっこう大変です。(第1章の最初のセクションだけで20項目もあります!)

まずは、もっと雑にやってみましょう!

- この資料では以下の3点を前提しています。
 - Git がインストールされていること git: 1.5 使い始める Gitのインストール
 - GitHub アカウントを持っていること GitHub Docs: GitHub でのアカウントの作成
 - **devtools** がインストールされていること <u>devtools: Tools to Make Developing R Packages Easier</u> install.pakcages("devtools")

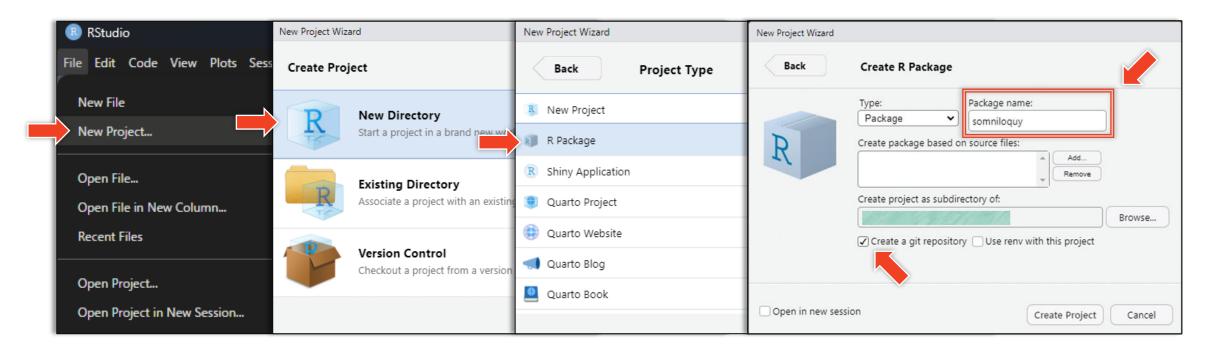


Getting started > 1. The Whole Game

- 1.1 Load devtools and friends
- 1.2 Toy package: regexcite
- 1.3 Preview the finished product
- 1.4 create_package()
- 1.5 use git()
- 1.6 Write the first function
- 1.7 use r()
- 1.8 load all()
- 1.9 check()
- 1.10 Edit DESCRIPTION
- 1.11 use_mit_license()
- 1.12 document()
- 1.13 check() again
- 1.14 install()
- 1.15 use_testthat()
- 1.16 use_package()
- 1.17 use_github()
- 1.18 use_readme_rmd()
- 1.19 The end: check() and install()
- 1.20 Review

1. パッケージのプロジェクトを作成する

- RStudio のメニューから File > New Project... > New Directory > R Package を選択し、パッケージ名とプロジェクトフォルダの置き場所を指定します。
 - ☑ Create a git repository の欄にチェックを入れておきましょう。
- このデモでは somniloquy という名前のパッケージを作ります。
 - 勉強会などで自作した雑多な関数を放り込むためのパッケージにすることを想定しています。



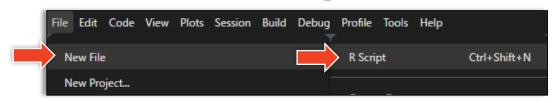
2. 作成されたフォルダを整理する

- パッケージと同名のプロジェクトフォルダが作成されたことを確認します。
- 自動的に開かれるスクリプト R/hello.R を閉じて、man/hello.Rd と一緒に削除します。

	パス (プロジェクトフォルダ内)	概要
	DESCRIPTION	パッケージのメタデータ(基本情報や依存関係など)を定義するファイル
	NAMESPACE	関数の外部へのエキスポートや外部からのインポートを管理するファイル
	R	パッケージの関数などを書いた R スクリプトの保存場所
	R/hello.R	削除 例示のための R スクリプト
	man	関数などのドキュメントに対応する Rd ファイルの保存場所
<i>!</i>	man/hello.Rd	削除 hello.R に対応する Rd ファイル
	.Rbuildignore	パッケージのビルド時に無視するファイルを定めるファイル
	.gitignore	Git のバージョン管理において無視するファイルを定めるファイル
	somniloquy(パッケージ名).Rproj	Rプロジェクトの設定ファイル

3.パッケージの関数を作成する

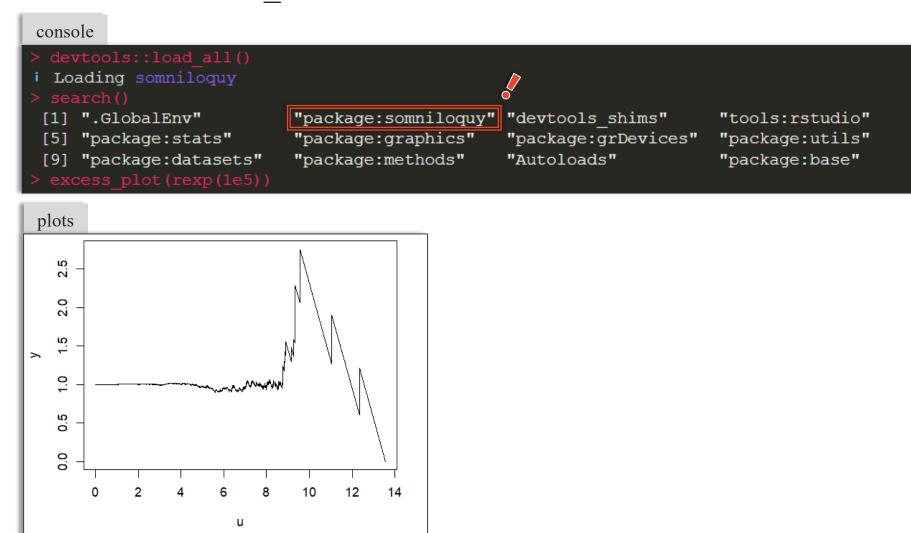
- File > New File > R Script で新しいファイルを作成し、File > Save as... で R/ に保存します。



- ファイル名は、そのファイルで定義する関数と同名にしておくと整理しやすいです。
- ファイルの作成と保存は、usethis::use_r("excess_plot")のようにしても OK。
- 作成されたスクリプトの中で関数を定義し、ファイルを上書き保存します。
 - ここでは平均/メディアン超過プロットを描くための関数 excess plot() を作成します。

4. 関数が正しく機能することを確認する

- devtools::load all()を使って作成した関数が正しく機能することを確かめます。



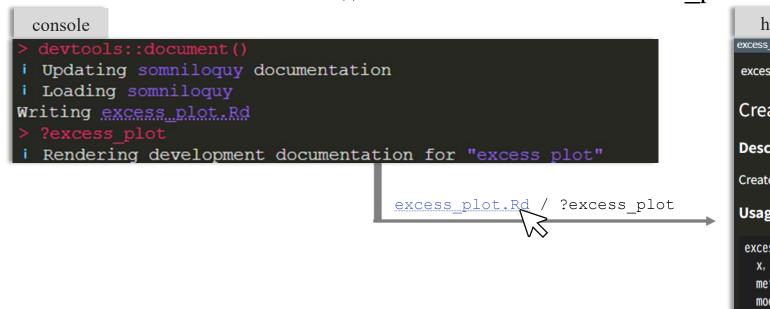
5. 関数がエクスポートされるように設定する

- 関数を作成した R スクリプトの冒頭に、その関数を端的に説明する「見出し」と @export タグを追加します。

```
excess_plot.R

1  #' Create Excess Plots
2  #'
3  #' @export
4  #'
5  excess_plot <- function(
```

- devtools::document() を実行して、man/excess plot.Rd を作成します。



```
help

excess_plot.Rd - Find in Topic

excess_plot {somniloquy} (preview) R Documentation

Create Excess Plots

Description

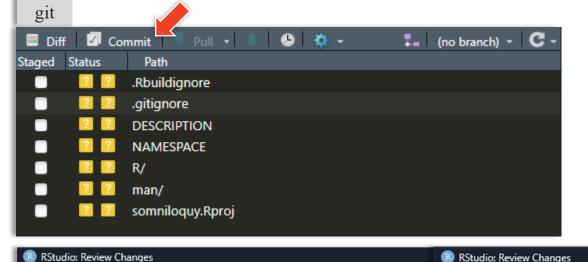
Create Excess Plots

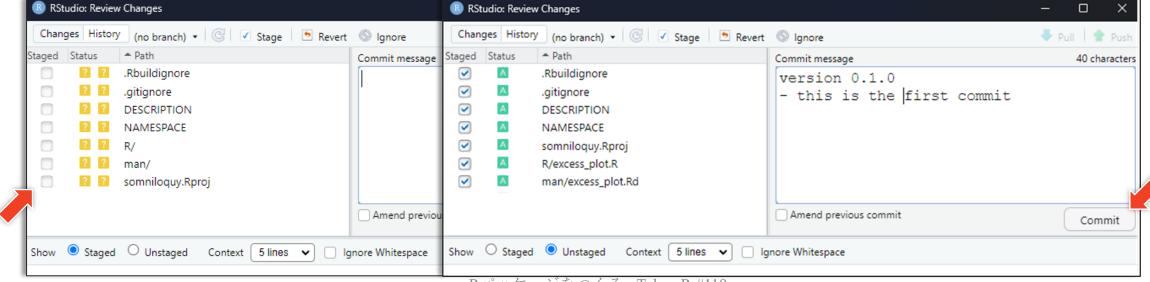
Usage

excess_plot(
    x,
    method = c("mean", "median"),
    mode = c("exact", "interpolate"),
    show_plot = TRUE,
```

6. すべての変更をコミットする

- R Studio の git タブにある ☑ Commit から進み、すべての変更をコミットします。





7. GitHub にアップロードする

- usethis::use github()でプロジェクトフォルダを GitHub にアップロードします。
 - 初めて use_github() を使うときは、ブラウザ経由で GitHub アカウントの認証を要求されます。

```
console

√ Creating GitHub repository "ryo-asashi/somniloquy".

√ Setting remote "origin" to

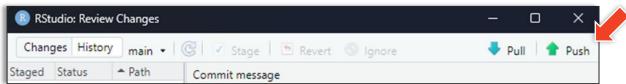
√ Adding "https://github.com/ryo-asashi/somniloguy" to URL.

✓ Adding "https://github.com/ryo-asashi/somniloguy/issues" to BugReports.
 i There is 1 uncommitted file:
! Is it ok to commit it?
1: Nope
2: Negative
3: For sure

√ Adding files.

✓ Making a commit with message "Add GitHub links to DESCRIPTION".
✓ Pushing "main" branch to GitHub and setting "origin/main" as upstream
  branch.
   Opening URL <a href="https://github.com/ryo-asashi/somniloguy">https://github.com/ryo-asashi/somniloguy</a>.
```

• 今後、ファイルの変更をコミットしたら、GitHub リポジトリにも変更をプッシュしましょう。



8. 完成! 👺 GitHub からインストールする

- これで完成! pak::pak() や devtools::install github() でのインストールが可能です。

```
console
→ Will install 1 package.
→ Will download 1 package with unknown size.
 somniloguy 0.1.0 [bld][cmp][dl] (GitHub: 123ef63)
i Getting 1 pkg with unknown size

√ Got somniloguy 0.1.0 (source) (1.56 MB)

✓ Downloaded 1 package (1.56 MB) in 941ms
i Packaging somniloguy 0.1.0

√ Packaged somniloguy 0.1.0 (4.3s)

Building somniloguy 0.1.0

√ Built somniloguy 0.1.0 (1.4s)

✓ Installed somniloguy 0.1.0 (github::ryo-asashi/somniloguy@123ef63) (119ms)

√ 1 pkg: added 1, dld 1 (NA B) [13.6s]

console
 library(somniloquy)
```

• なお、開発上は GitHub を経由せず devtools::install() でインストールしても問題ありません。

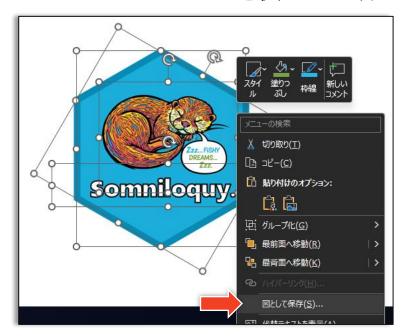
Optional. DESCRIPTION を更新する

- DESCRIPTION は、パッケージのメタデータ(基本情報や依存先パッケージ)を記載するファイルです。ここでは、Title / Authors@R / Description などを書き換えておきます。
 - このうち、Title は GitHub リポジトリの About 欄にも反映させておくとよさそうです。

```
DESCRIPTION
   Package: somniloquy
 2 Type: Package
   Title: A Collection of Miscellaneous R Functions
   Version: 0.1.0
   Authors@R: c(
       person (
          "Ryoichi", "Asashiba",
         email = "ryoichi.asashiba@gmail.com",
         role = c("aut", "cre")
11
    Description: This package is a personal collection of miscellaneous R functions
   License: What license is it under?
   Encoding: UTF-8
   LazyData: true
16 RoxygenNote: 7.3.2
17 URL: https://github.com/ryo-asashi/somniloguy
18 BugReports: https://github.com/ryo-asashi/somniloguy/issues
```

Optional. パッケージロゴを作成する

- PowerPointで外枠・背景・タイトルを用意します。
 - 外枠と背景はそれぞれ「挿入>図形>六角形」で作成 「図形の書式」でサイズを 10.22 cm×11.82 cm に変更して 90 度回転
 - 外枠のみ、「図形の枠線>太さ」で枠の幅を 12pt に設定
- ロゴ用のイラストは生成AIに描いてもらいました。
- 作成したロゴは man/figures/logo.png に保存します。
 - パーツをまとめて選択して「右クリック>図として保存」





Optional. README を追加する

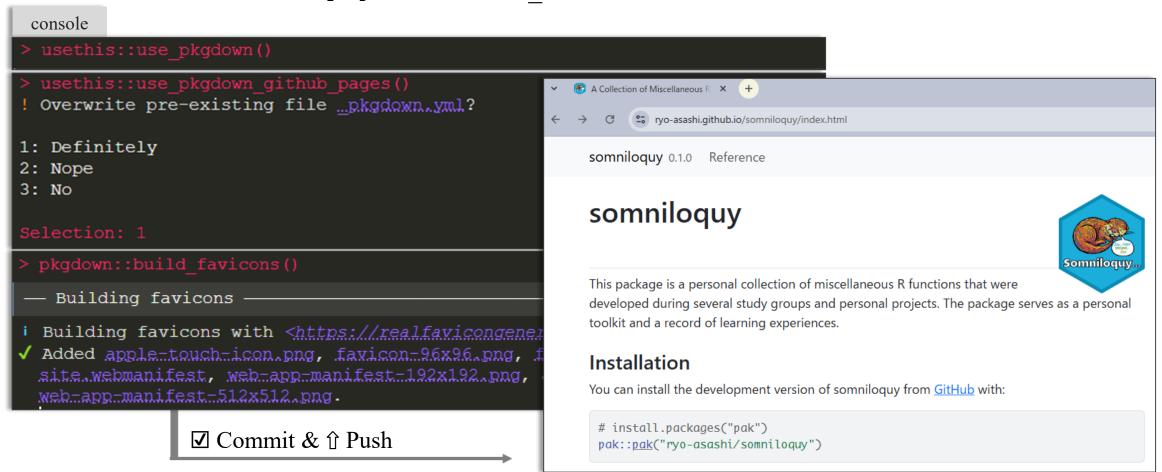
- usethis::use readme rmd() を実行して README.Rmd を作成します。

- README.Rmd の16行目を以下のように書き換えるとロゴを入れられます。
- # somniloquy
- 21行目は DESCRIPTION の Description をコピーしておけば OK です!
- 32行目以下は「パッケージの使用例」のサンプルです。とりあえず消しておきましょう。
- Knit を実行して README.md を出力します。



Optional. GitHub Pages で Webサイトを作る

- usethis::use_pkgdown() と usethis::use_pkgdown_github_pages() でパッケージを紹介するための Web サイト (GitHub Pages) を作れます。
 - ロゴを作ってあるなら、pkgdown::build favicons() でファビコンも作成できます。



この資料で説明しなかったこと (の一部)

- 他のパッケージの関数を利用する(baseを除く)
 - DESCRIPTION の Imports: か Suggests: にパッケージ名を記載し、スクリプト内では stats::median() のようにパッケージ環境を指定して参照します。usethis::use package("stats") でも OK。
 - 特に、ggplot2 を利用する場合は vignette ("ggplot2-in-packages") が参考になります。
- パッケージのライセンスを設定する
 - ライセンスは、パッケージのコードを著作権者でない人がどのように利用できるかを規定します。
 - usethis::use mit license() を実行すると MIT ライセンスを設定したものとして体裁が整います。
- パッケージの関数に対するコードテストを実行する
 - usethis::use testthat() によって testthat パッケージのコードテスト機能を利用できます。
 - その後は usethis::use_test() 関数などを用いてテスト用スクリプトを追加していきます。
- パッケージの問題点をチェックする
 - devtools::check() を用いてパッケージが問題を抱えていないかを包括的にチェックできます。
 - 潜在的な問題があれば、errors, warning, notes という 3 種類の注意が表示されます。 (CRAN への投稿を考えていないなら、 errors 以外はそれほど深刻にとらえなくてもよいと思っています。)

CRANでの公開までの苦しい道のり

- 第一の壁:devtools::check()
 - 気軽にいつものようにコードを書くと、check() に色々なことを指摘されます (外部パッケージの非エクスポート関数を使うと警告、ggplot や do.call 関係でexpression をそのまま使うと警告...などなど)。
 - 慎重に書き直せば大抵は回避できますが、全部対処するのはそれなりに骨が折れました...。
- 第二の壁: <u>Davis Vaughan/extrachecks</u>
 - 先人たちが、CRANから受けたアドホックな指摘を受けた経験をリストアップして共有してくれています。

extrachecks

This is a place to dump extra ad-hoc checks that CRAN does that are not checked for by devtools::check(). Some of them are generally useful, some of them are highly specific, but all of them are reasons that an R package has been rejected by CRAN. The hope is that by making them public, we can lower the number of rejections by making package developers more informed.

First release:

- ☑ usethis::use cran comments()
- ☑ Update (aspirational) install instructions in README
- ☑ Proofread Title: and Description:
- ☑ Check that all exported functions have @return and @examples
- ☑ Check that Authors@R: includes a copyright holder (role 'cph')
- ☑ Check licensing of included files
- ☑ Review https://github.com/DavisVaughan/extrachecks

Prepare for release:

- ☑ git pull
- ☑ urlchecker::url_check()
- ☑ devtools::build readme()
- ☑ devtools::check(remote = TRUE, manual = TRUE)
- ☑ devtools::check win devel()
- ☑ git push

Submit to CRAN:

- ☑ usethis::use_version('minor'): -> incremented to 0.5.0
- ☑ devtools::submit_cran()
- ☑ Approve email

Wait for CRAN...

- ☑ Accepted 🎉
- ☑ usethis::use github release()
- ☑ usethis::use dev version(push = TRUE)

最後にして最大の壁:References

- 第三の壁:<u>The CRAN Cookbook</u> References
 - 理想としては、各パッケージの説明文には、ユーザーがパッケージの背景理論を探求するのに役立つ参考文献を少なくとも1つ含めるべきです。文献は「GA」パッケージのような形式で記述してください。Ideally, each package should include at least one reference in the description text to help users further explore the theory behind the package. These references should be formatted as shown in the 'GA' package:



• 一応、Reference は必須ではないようですが、今回はパッケージの実装を説明する英語の論文を準備して arXiv [2506.08338] に投稿しました(これでだいたい半年くらいパッケージの公開が遅れました)。

midr: Learning from Black-Box Models by Maximum Interpretation Decomposition

by Ryoichi Asashiba, Reiji Kozuma, and Hirokazu Iwasawa

Abstract The use of appropriate methods of Interpretable Machine Learning (IML) and eXplainable Artificial Intelligence (XAI) is essential for adopting black-box predictive models in fields where model and prediction explainability is required. As a novel tool for interpreting black-box models, we introduce the R package midr, which implements Maximum Interpretation Decomposition (MID). MID is a functional decomposition approach that derives a low-order additive representation of a black-box model by minimizing the squared error between the model's prediction function and this additive representation. midr enables learning from black-box models by constructing a global surrogate model with advanced analytical capabilities. After reviewing related work and the theoretical foundation of MID, we demonstrate the package's usage and discuss some of its key features.

最後に少しだけ...

- midr パッケージの宣伝をさせてください!
- カラーテーマ ("color.theme") 機能はもしかしたら便利なことがあるかも?
 - midr パッケージには、外部パッケージへの依存をできるだけ少なく抑えるという思想から、可視化に とって重要な"色"の取り扱いを自前で実装しています。
 - 以下のシンタックスに従う文字列で、カラーテーマをカスタマイズして呼び出すことができます。

"[(source)/](name)[r][@(type)][?(query)]"

color.theme("tokyo")

Sequential Color Theme: "tokyo"

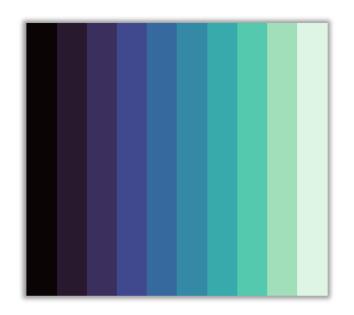


Sequential Color Theme: "tokyo"

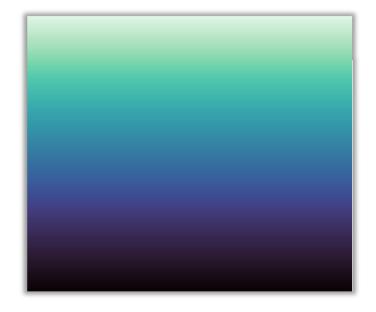
カラーテーマは二つのメソッドを持ちます

- カラーテーマ(環境)は …\$palette() と …\$ramp() という関数を持ちます。 それぞれ、正の整数 n と [0,1] 区間に値を持つベクトル x に対して色名を返します。

```
n <- 10
barplot(height = rep(1, n),
       space = 0, border = NA, axes = FALSE, axes = FALSE,
```



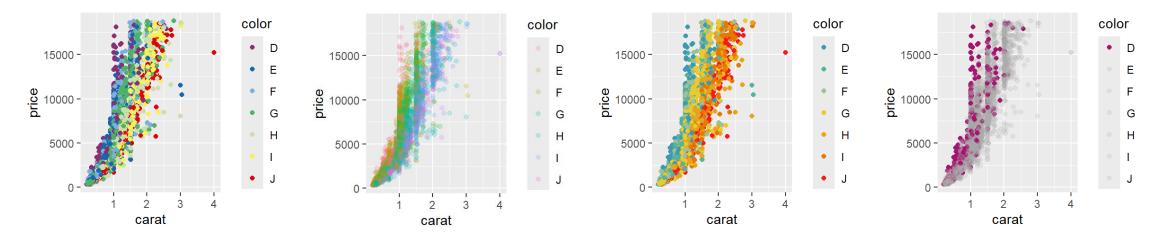
```
x < - seq(0, 1, length.out = 100)
                                     image(z = t(1:100),
col = color.theme("mako")palette(n)) col = color.theme("mako")$ramp(x))
```



ggplot用のインターフェースもあります

- ggplot 用に scale_color_theme() と scale_fill_theme() を実装しています。
 - ・ テーマは最大 234 種類のプリセットが用意されており、さらに、テーマを自作することもできます。

```
d <- diamonds[sample(nrow(diamonds), 1e4),]
p <- ggplot(d) + geom_point(aes(carat, price, col = color))
p + scale_color_theme("discreterainbow")
p + scale_color_theme("HCL?alpha=.2")
p + scale_color_theme("Zissou 1@q")
p + scale_color_theme("highlight?base='#aaaaaa30'")</pre>
```



ggplot用のインターフェースもあります

- ggplot 用に scale_color_theme() と scale_fill_theme() を実装しています。
 - テーマは最大 234 種類のプリセットが用意されており、さらに、テーマを自作することもできます。

```
v <- ggplot(faithfuld) + geom_tile(aes(waiting, eruptions, fill = density))
set.color.theme(c("#003f5c", "#7a5195", "#ef5675", "#ffa600"))
v + scale_fill_theme("viridis")
v + scale_fill_theme("viridis_r")
v + scale_fill_theme("newtheme_r?alpha=.7")
v + scale_fill_theme("bukavu", middle = 0.017)</pre>
```

