

# RとQuartoではじめるデータサイエンス《2026》

## #7 テーブル

苅谷千尋

Wed, 27, May, 2026

### 目次

1. 前回の振り返り
2. テーブル
3. 統計分析
4. レポート課題の準備・相談

### 0. 本日の目標

1. データをcsvで書き出せるようになる
2. 集計表・クロス集計表を作ってみる
3. 検定・回帰分析に触れてみる

### 2. テーブル

#### (1) データの書き出し

- `write_csv()` 関数
- データフレームや計算結果を代入して、`write_csv()` に渡す

#### (2) テーブル出力関数

- 複数のパッケージ・関数が乱立
  - `gt()` : 得意な出力形態 : HTML
    - 公式サイト ; `install.packages("gt") / library(gt)`
  - `kable()` : 得意な出力形態 : HTML / Word / PDF / pptx 全般
    - 公式サイト ; `install.packages("knitr") / library(knitr)`
  - `flextable()` : 得意な出力形態 : Word・pptx
    - 公式サイト ; `install.packages("flextable") / library(flextable)`
  - `tt()` : 得意な出力形態 : PDF
    - 公式サイト ; `install.packages("tinytable") / library(tinytable)`

#### (3) 集計表

- `tabyl()` 関数 (公式サイト)
  - 公式サイト ; `install.packages("janitor") / library(janitor)`
  - データをきれいに整理するためのRパッケージ

## i) 1変数集計

```
tabyl(species) |> # 件数表の作成
```

## ii) 2変数集計=クロス集計表

```
tabyl(species, island) |> # 2変数を指定
```

## 3. 統計分析

### (1) カイ二乗検定

- クロス集計表を作り、変数間の値の違いが偶然かどうかを調べる
- 例：ペンギンの種類と島に関係はあるのか

```
chisq.test()
```

- `tbl_summary()` 関数
  - 公式サイト ; `install.packages("gtsummary") / library(gtsummary)`

### (2) 回帰分析

#### (1) 単回帰分析

- 被説明変数：～の左側；説明変数：～の右側
  - 例：くちばしが長いペンギンほど体重も重いかを調べる

```
lm(body_mass ~ bill_len, data = _) # 被説明変数、説明変数を指定。data = _は「直前のデータを使う」の意
```

#### (2) 重回帰分析

- 複数の説明変数を挙げ、被説明変数に対して、より説明力のある変数を特定する
- 方法：単回帰分析のコードに、説明変数を `+` を付けて足す

## 宿題

### 授業の感想

- 回答先：Google Forms
- 締め切り：5月29日（金）23時59分

### レポート課題

- ファイル：ZIPファイル
  - ファイル名：どこかに氏名を特定できる文字を入れて下さい
  - 例：kariya.zip
- 提出先：dropbox
- 締め切り：6月3日（水）10時30分