

第1章: イントロダクション

今井耕介 著

『社会科学のためのデータ分析入門 (QSS)』

2026-03-09

1.1 R 入門

計算操作 (Arithmetic Operations)

▶ Rは強力な計算機として使用できる。

```
# 基本的な四則演算
```

```
5 + 3
```

```
## [1] 8
```

```
5 - 3
```

```
## [1] 2
```

```
5 / 3
```

```
## [1] 1.666667
```

```
# べき乗と関数の使用
```

```
5 ^ 3
```

```
## [1] 125
```

1.2 オブジェクト (Objects)

オブジェクトの作成と上書き

- ▶ `<-` を用いて、計算結果やデータを名前付きの「オブジェクト」として保存する。

```
result <- 5 + 3  
result
```

```
## [1] 8
```

```
# 文字列の保存
```

```
kosuke <- "instructor"  
kosuke
```

```
## [1] "instructor"
```

- ▶ オブジェクトは後から上書きできる。また、大文字と小文字は区別される (`result` と `Result` は別物)。

1.3 ベクトル (Vectors)

ベクトルの作成

- ▶ `c()` 関数 (concatenate) を用いて、複数の要素をひとまとめにした「ベクトル」を作成する。

```
# 世界人口データ (千人単位)
```

```
world.pop <- c(2525779, 3026003, 3691173, 4449049,  
              5320817, 6127700, 6916183)
```

```
world.pop
```

```
## [1] 2525779 3026003 3691173 4449049 5320817 6127700 6916183
```

要素の抽出

▶ [] を用いてベクトルの一部を取り出すことができる。

```
# 2番目の要素を取り出す
```

```
world.pop[2]
```

```
## [1] 3026003
```

```
# 2番目と4番目を取り出す
```

```
world.pop[c(2, 4)]
```

```
## [1] 3026003 4449049
```

```
# 3番目以外のすべての要素を取り出す
```

```
world.pop[-3]
```

```
## [1] 2525779 3026003 4449049 5320817 6127700 6916183
```

1.4 関数 (Functions)

組み込み関数の使用

▶ Rには多くの関数が用意されている。

```
# 長さ、最小値、最大値、平均
```

```
length(world.pop)
```

```
## [1] 7
```

```
min(world.pop)
```

```
## [1] 2525779
```

```
max(world.pop)
```

```
## [1] 6916183
```

```
mean(world.pop)
```

```
## [1] 4579529
```

カスタム関数の作成

- ▶ 自分で独自の関数を定義することもできる。

独自の要約関数

```
my.summary <- function(x){  
  s.out <- sum(x)  
  l.out <- length(x)  
  m.out <- s.out / l.out  
  out <- c(s.out, l.out, m.out)  
  names(out) <- c("sum", "length", "mean")  
  return(out)  
}
```

```
my.summary(world.pop)
```

```
##          sum    length    mean  
## 32056704         7 4579529
```

1.5 データファイル (Data Files)

CSV ファイルの読み込み (1)

- ▶ `read.csv()` 関数を用いて、外部のデータセットを R に読み込む。
- ▶ 原則として、データファイル (`UNpop.csv` など) を作業ディレクトリに保存して読み込むことが推奨される。

1. ローカルに保存した `UNpop` データの読み込み (推奨)

```
UNpop <- read.csv("UNpop.csv")
```

(参考) URL から直接読み込むことも可能

```
# UNpop <- read.csv("https://ayumu-tanaka.github.io/QSS/QSS_Data/UNpop.csv")
```

データフレームの次元 (行数・列数) を確認

```
dim(UNpop)
```

```
## [1] 7 2
```

CSV ファイルの読み込み (2)

▶ 読み込んだデータの中身を `head()` で確認します。

```
# 最初の 3 行を表示  
head(UNpop, n = 3)
```

```
##   year world.pop  
## 1 1950   2525779  
## 2 1960   3026003  
## 3 1970   3691173
```

データフレームからの変数の抽出

- ▶ \$ 演算子を用いて、特定の列 (変数) にアクセスする。

```
# world.pop 変数の抽出
```

```
UNpop$world.pop
```

```
## [1] 2525779 3026003 3691173 4449049 5320817 6127700 6916
```

- ▶ 特定の行や列にインデックス [行, 列] でアクセスすることも可能。

```
# 1~3行目のすべての列
```

```
UNpop[1:3, ]
```

```
##   year world.pop
```

```
## 1 1950    2525779
```

```
## 2 1960    3026003
```

```
## 3 1970    3691173
```

1.6 パッケージ (Packages)

R パッケージの利用

- ▶ パッケージをインストールすることで、R の機能を拡張できる。
- ▶ 例: `foreign` パッケージを用いて、Stata (`.dta`) や SPSS (`.sav`) などの他ソフトウェアのデータを読み書きする。

```
# インストール (初回のみ)  
install.packages("foreign")
```

```
# パッケージの読み込み  
library(foreign)
```

1.7 まとめ

第1章のまとめ

- ▶ **R と RStudio** の基本的な操作方法を習得した。
- ▶ **オブジェクトとベクトル**を用いたデータの保存と操作方法を学んだ。
- ▶ **組み込み関数や自作関数**を用いて、データを要約する方法を理解した。
- ▶ **read.csv()** を用いてデータフレームを読み込み、操作する方法を学んだ。
- ▶ データ分析に必要な基礎的なコーディングスキルを身につけた。