

「ミクロ政治データ分析実習」第5回課題

情21-0170 関大太郎 ←自分の学籍番号と氏名
作成日: 2023年01月24日

Rとは何か ←最も大きい見出し (#) 太字↓ イタリック↓

Rは統計、データ分析、作図のためのインタープリタープログラミング言語である。Rという名前は二人の開発者Ross Ihaka、Robert Clifford Gentlemanのイニシャルに由来する。R以外にも、統計・データ分析のために利用可能なソフトウェアはたくさんある。無料でありながら優れたソフトウェアとして以下のような選択肢もある。

ソフト・言語名	備考
PSPP	SPSSにととてもよく似た無料ソフトウェア
JASP/jamovi	裏で動いているのはR
gretl	時系列分析など、計量経済学で利用される手法に特化したソフト
GNU Octave	MATLAB とほぼ同じ文法をもつ無料言語
HAD	Excelベースのデータ分析マクロ

このように世の中には様々な統計・データ分析・作図のためのソフトウェアが存在するが、なぜRを使うだろうか。

Rの長所と短所 ←最も大きい見出し (#)

Rの長所 ←二番目に大きい見出し (##)

- Rは無料で、オープンソースで、多くのプラットフォーム（訳注: macOS, Linux, Windowsなど）で利用できる。
- オンライン、オフラインの両方で、多様なRコミュニティがある。
- 統計モデリング、機械学習、可視化、データ読み込みおよびハンドリングのための膨大なパッケージが用意されている。
- 分析結果を伝達する強力なツールを提供している。
- 代表的な統合開発環境であるRStudioはデータサイエンス、対話型のデータ分析、そして統計的プログラミングが必要とするものに最適化されている。
- 多くの統計学や機械学習の研究者は自分の研究成果とRパッケージを同時に公開している。
- データ分析を根強くサポートする言語である。
- Rはデータサイエンスに非常に有効である関数型プログラミングのための最適な環境を提供している。
- RはC、C++、Fortranのようなハイパフォーマンス言語と容易に結合できるように設計されている。
- 宋さんの激推言語である。 ←取り消し線

Rの短所 ←二番目に大きい見出し (##)

一見、完璧にも見えるRだが、以下のようなデメリットも存在する。

- 多くのRコードは「今の」問題を解決するために迅速に書かれたものであり、あまりエレガントでも、速くも、読みやすくもない。
- 開発されたパッケージは、R内蔵のパッケージさえも一貫性が乏しい。
- Rは格別に速い言語ではない。

RMarkdownの例 ←最も大きい見出し (#)

RMarkdownを使うと、以下のようにコードと分析結果を一つの文書としてまとめることができる。

356 * 356 チャンク化すること

[1] 126736 リンク先を指定
(<https://www.jaysong.net/RBook/rmarkdown.html>) → 教科書を参照すること。

```
library(tidyverse)
iris %>%
  ggplot() +
  geom_point(aes(x = Sepal.Length, y = Sepal.Width, color = Species))
```

 チャンク化すること