

「ミクロ政治データ分析実習」第5回課題

作者

情21-0170 関大太郎

公開

2023年3月16日

Rとは何か

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にとっ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

Rの長所

1. 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。
2. 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。
3. 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。
4. 分析結果を伝達する強力なツールを提供している。
5. 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。
6. 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。
7. データ分析を根強くサポートする言語である。
8. 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。
9. 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。
10. 宋さんの激推言語である。

Rの短所

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

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

[1] 126736

計算結果だけでなく、図まで文書に埋め込むことができる。RMarkdownの詳細については **教科書** を参照すること。

