

Amazon AWS を用いた RStudio Server 環境構築の必要性 に関する予備的な報告

宇都宮 譲¹⁾

1) 長崎大学 経済学部

yuzuru@nagasaki-u.ac.jp

Is it necessary to set up RStudio Server environment using Amazon AWS? A preliminary report

Yuzuru Utsunomiya

1) Faculty of Economics, Nagasaki University

概要

本報告は、筆者が統計解析ハンズオンセミナーを実施した経験から、RStudio Server 導入するに至った経緯を報告することを目的とする。インストールの困難とマシンスペックがばらついてハンズオンセミナー実施を困難とした。結果、サーバアプリ導入と Amazon AWS サーバ利用による確実な実施環境構築が必要であるという結論を得た。

1. 目的

本報告は、Amazon AWS 上に RStudio Server 環境を構築するに至った背景紹介を目的とする。

筆者は JICA (独立行政法人国際協力機構) 調査団員として、ベトナム社会主義共和国 (ベトナム) へ統計解析環境 R についてハンズオンセミナーを実施した。実施時期は 2018 年 1 月および 2018 年 12 月である。研修は、同国において技術協力および円借款の枠組にて実施される大学高度化プログラムの一環にて実施された。

研修を実施するに至った理由は、以下に示すとおりである。第一、目的に応じた統計モデルを構築可能なソフトウェアに習熟する必要があることである。同プログラムに参加する研究者はさまざまな学術的背景を有する。取り扱うデータも遺伝解析から社会調査、水質データまで多岐にわたる。これらを組み合わせて研究を進めようとする場合、従前にはなかった統計モデルとなる公算大である。かような場合、柔軟に統計モデルを組み立てられる環境が望ましいことから、筆者は R を統計解析環境としてふさわしいという結論を得た。R は無料統計解析環境である [1]。科学の世界において事実上の標準といえる地位を占めるに至った。無料で柔軟なモデリングが可能であり、自習用教材も多数出版されている (たとえば [2][3])。R を使用する場合、RStudio と呼ばれる IDE から用いるこ

とが簡便である。

第二、海賊版を用いる誘引を減殺することである。同地においては、さまざまな海賊版ソフトウェアがきわめて安価にて販売される。学部学生が高価なフォトタッチソフトや三次元 CAD をインストールしたノート PC を所持している様子を頻繁に見かける。かようなソフトウェアは、街角にて数百円程度にて販売される。統計解析については、SPSS や SAS、STATA といったソフトウェアが流通しているようである。研究倫理という観点から、かような状況は明らかによりしくない。R を用いるならば、海賊版を使う必要が薄れる。

2. 対象と方法

研修会期間・内容・方法は以下の通りである。研修期間は 5 日間、朝 8 時から 11 時までおよび 13 時から 16 時までである。内容は簡単な四則演算、データハンドリング、折れ線グラフ作図、記述統計量算出、統計的仮説検定、および回帰分析である。参加者は当初 10 名程度を予定したが、結局は 20 名となった。初日には R と RStudio をインストールするために、個人所有 PC を持参してもらった。個人所有 PC がないもしくは壊れている場合に備えて、大学所有 PC を準備した。大学所有 PC には、あらかじめ R と RStudio をインストールした。

3. 結果

本ハンズオンセミナーにおける最大の難関は、R ないし RStudio をインストールできない場合が散見されたことである。コンピュータとくにキーボードに対する知識・経験不足、CUI に対する苦手意識、英語による意思疎通困難、統計学に関する基礎知識不足なども困難であった。しかし、インストールできないことが最大の困難である。講習会自体を実施困難にするからである。

R をインストールしようとする、さまざまなエラーにてインストールできない事態が散発した。たとえば、2018 年 1 月に実施した最初の講習会においては、参加者 23 名中 5 台においてインストールが成功しなかった。概して M 社製タブレット PC にインストールできないが多かったように感じられる。同様な事態は、筆者が勤務校において R を用いる場合にも発生したことから、国・地域特有の事情とはいえないようである。かような方には用意した大学所有 PC を用いて講習会を実施した。ただし、大学所有 PC についても、まれに R をインストールできないことがあった。

R をインストールできないことが予想しないデメリットをもたらす。講習会に参加した受講証が効力を有する社会において、講習会に参加できないことはユーザの社会的評価を下げる可能性がある。すなわち、インストールに失敗することがその後の人生にわずかな影響を与える。R を使用する機会を失うことで逸失する便益も計り知れない。近年とくに R における機能拡充は著しく、データ分析から論文執筆まで研究活動を一通り R と RStudio のみにて実施することもできるようになった。さまざまなソフトウェアを切り替えながら使う不便が解消されたのである。

マシンスペックがばらつくことも困難であった。同地教職員が保有する PC の能力は、想像を超えて大きくばらつく。Windows 10 が快適に動作するノート PC を保有する教職員もいれば、Windows XP が購入時にはインストールされていたようなマシンを使う教職員もいる。なかにはキートップが外れているマシンを用いる教職員さえいる。R においては、マシンスペック差が特に作図・描画において大きな計算時間差を生む。結果として新しい PC をもっているユーザは退屈な時間を過ごすことになる。

4. 考察

筆者が実施した R ハンズオンセミナーにおいて、PC にインストールする困難とマシンスペックばらつきが大きな問題であることを発見した。以上から、R ハンズオンセミナー実施について、筆者は以下に示す結論を得た。

第一、サーバアプリを導入することが望ましい。R はサーバ上においても稼働する。RStudio には RStudio Server と呼ばれるサーバ製品が存在する。機能は限定されるが無料版も存在する。実用上はまったく差し支えない。サーバアプリを用いるメリットは、PC のスペックがばらつくことを勘案する必要がないことである。ただしサーバが随時更新される必要がある。サーバアプリを使うための端末に使う PC をネットワークに接続する必要があることもデメリットに見える。しかしながら同国はネットワークに接続する環境に不自由することはない。Wi-Fi 環境を見つけることにさほどの苦労はない。

第二、Amazon AWS が望ましい。Amazon AWS は、Amazon 社が提供するクラウドコンピューティング環境である。サーバ機は当然ながら随時更新される。同サービスは R や RStudio Server 利用が可能である。エデュケーションプランを用いると、75USD/人・年の無料利用枠が付帯する。これは学部学生や R を使う頻度が低い方には十分すぎる利用枠である。R ハンズオンセミナーには好適である。

現在、2019 年 10 月稼働を目指して環境構築中である。構築完了後に発生した事例報告は今後の課題としたい。

参考文献

- [1] R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- [2] Teetor, P. (2011). R Cookbook. O'Reilly,
- [3] Taveras, J. (2016). R for Excel Users: Introduction to R for Excel Analysts. CreateSpace Independent Publishing.