

再現可能性を重視したオンラインでのジオコンピューテーション教育

堤田成政*

Teaching reproducible Geocomputation Online

Narumasa Tsutsumida*

The COVID-19 crisis of 2020 has led us to the delivery of online lectures for geographic information (GI) in the undergraduate/postgraduate education all over the world to avoid infection.

Under this circumstance, educators have struggled with teaching GI theories and tools effectively online. This paper demonstrates my experience of the delivery of a lecture for geographic information science to postgraduate students in the Department of Information and Computer sciences, Saitama University, Japan.

In this lecture, a reproducibility of geocomputational research/education was explored by using multiple applications including Git/Github, Docker, R, and Rstudio server built in a Docker container.

Students learnt GI theories and geocomputation skills by R, and submitted their own report with maps of their-defined themes by this reproducible approach.

Keywords: R (R) , Github classroom (Github classroom) , Docker (Docker)

1. はじめに

2020年のCOVID-19の危機を受けて、感染拡大を防ぐため世界中で高等教育の授業がオンラインで展開された。グラフィカルユーザーインターフェース(GUI)ツールを多用する地理情報に関する授業も例外なくオンライン化し、世界中で関連の課題が議論されてきた(例えば Global GIScience Education, <https://www.globalgisscienceeducation.org/>)。本稿では、そのような状況下において実施した授業の経験をまとめたものである。コロナ禍以前より叫ばれてきた「再現可能な研究」(Peng, 2011)を重要視し、GUIツールの活用が主流である地理情報教育における新たなアプローチを試みている。

筆者は2021年前期に埼玉大学大学院理工学研究科で20名の受講生(博士前期課程学生15名、学部生5名)が受講した地理情報科学特論(15回、2単位)を担当した。コロナ禍において感染状況が安定

していなかったことから全ての講義をオンライン形式で実施することとした。対面講義では学生からのちょっとした質問や相談を受けやすい一方で、オンライン授業ではそのようなコミュニケーションが比較的難しい。さらに、コンピュータを用いた実習を伴う授業では各受講生のコンピュータ環境が異なるため、一つの課題を実施するにしてもアプリケーションのインストールや操作方法などが異なってしまう、それらの対応をオンライン講義で実施することは困難である。そのため、本講義ではOS環境に依存しない、再現性の極めて高い演習・分析環境を整備し、受講生と共有することで上記の課題を克服した。本稿では再現可能性を重視したこれらの経験を紹介する。

2. 授業概要

地理情報科学特論は、オープンソースの統計解析

* 正会員 埼玉大学大学院理工学研究科 (Graduate school of Science & Engineering, Saitama University)
〒338-8570 埼玉県さいたま市桜区下大久保255 E-mail : narut@mail.saitama-u.ac.jp

ソフトウェア R (R Core Team, 2021) による地理情報の処理・分析技術を習得することを目的とした。R はプログラミング言語の一つであり、ArcGIS (Environmental Systems Research Institute, 2011) や QGIS (QGIS Development Team, 2009) などの GUI ベースのソフトウェアと異なり、CLI (Command Line Interface) により操作する。R には地理情報分析のための様々なパッケージが公開されており、地理情報を扱うツールとして汎用性の高いアプリケーションとなっている。GUI の地理情報システム (GIS) ではデータが加工されるごとに結果を容易に視覚化 (地図化) できるため、地理情報処理のプロセスの理解が比較的容易である。CLI ベースの分析はそのような利点はないものの、一連の分析工程がプログラムとして保存されるため、実行環境が再現できるのであれば誰でも同一のプログラムを実行し、同じ結果を得ることができる利点がある。さらに、近年では地理情報がビッグデータ化し、大量かつ多様な種類の地理情報データが入手しやすくなってきた点を鑑みると、CLI を用いた分析手法は優位性があるといえるだろう。このように、地理情報データ分析の計算機科学的アプローチであるジオコンピューテーションと呼ばれる分野が発展してきた (Abrahart et al., 2000; Brunsdon and Singleton, 2015; Longley et al., 1998)。

以上の背景を踏まえ、授業で扱う内容と実施回数には以下の通りとした。

- I. 地理情報科学概論 (2回)
- II. 再現可能な研究 (1回)
- III. R 入門 (4回)
- IV. R によるジオコンピューテーション (5回)
- V. 演習 (3回)

I, II に関しては一方向型の講義, III, IV, V については演習を伴う実践的な内容とした。III, IV はそれぞれ An Introduction to R (Douglas et al., 2021), Geocomputation with R (Lovelace et al., 2019) というオンライン上で公開されている教科書を用いた。これら 2つの教科書は R の bookdown (Xie, 2016) パッケージで作成された html 形式の教科書であり、それぞれ CC BY-NC 4.0, CC BY-NC-ND 4.0 ライセン

スで公開されている。そのため、受講生は教科書を購入することなく、かつデバイスにとらわれずインターネットに接続することさえできればどこでも学習することが可能となる。

V は, I~IV で学習した内容を踏まえ、自由テーマのレポートを課した。レポートは後述する演習環境にて, R Markdown (Xie et al., 2020) により記述した再現可能な (コードを含めた) html 形式のファイルを指定した。

3. 環境構築

オンライン講義という制約と再現可能な研究環境を受講生に提供するため、図 1 のように複数のアプリケーションを活用した。zoom (<https://zoom.us>) は所属大学が契約している授業や会議のためのビデオ会議システムであり、地理情報科学特論ではリアルタイム方式で実施した。しかし、コミュニケーションツールとしての zoom のチャット機能は講義中のみしか対応できず不十分と感じたため、slack (<https://slack.com>) のチャンネルを開設し、そちらで関連連絡や質問を受け付けるようにした。これにより、全体アナウンスや個別の質問・相談対応のみならず、「#Questions」というチャンネルを通じて QA を受講生全員で共有することにより双方向のコミュニケーションを図った。授業で必要な実行環境構成ファイルやプログラムは Github Classroom (<https://classroom.github.com>) を通じて提供した。Github Classroom は、Github のプラットフォームを拡張・応用し、プログラミング演習をサポートする仕組みである。教員側は教材であるレポジトリを提供し、受講生は学籍番号と Github アカウントを紐付け、提供されたレポジトリを fork^{*1} する。受講生はそのレポジトリから環境を構築し、ファイルを追加・編集・更新し、commit^{*2} したうえで Github に push^{*3} することで教員側に課題を提出することができる。本講義では、地理空間関連パッケージが内包された R 環境と Rstudio server (<https://www.rstudio.com/products/rstudio/#rstudio-server>) を構成する Docker image^{*4} (<https://hub.docker.com/r/rocker/geospatial>) と、Geocomputation with R が公開しているレポジトリ (<https://github.com/Robinlovelace/geocompr#geocomputation-with-r>) を基盤とする独自の演習用レポジト

「2021GISClass_docker_renv」(Tsutsumida, 2021a) を作成し、提供した。

受講生は、fork したレポジトリから Docker コンテナを構築し課題に取り組むことができるほか、Geocomputation with R の教科書の内容をハンズオン形式で学習することも可能である。これらを実現するアプリケーションが Docker

(<https://www.docker.com>) である。Docker は、Docker engine を OS にインストールすることでコンテナ*5 を構築することができ、実行している OS に依存しない仮想環境を作り上げることができる。仮想コンテナ環境の構成は Dockerfile に記載されており、本節で説明したように Github などを通じて共有することができる。受講生は、このように Github Classroom を通じて fork した

「2021GISClass_docker_renv」レポジトリ内の Dockerfile をもとに、Docker コンテナを構築する(図 2)。

コンテナを起動すると Rstudio server が立ち上がり、任意の web ブラウザでアクセスすることで、関連パッケージがインストールされた R の実行環境が出来上がる。このような工夫により、受講生の PC に依存しない実習環境を提供する事が可能となる*6,7。

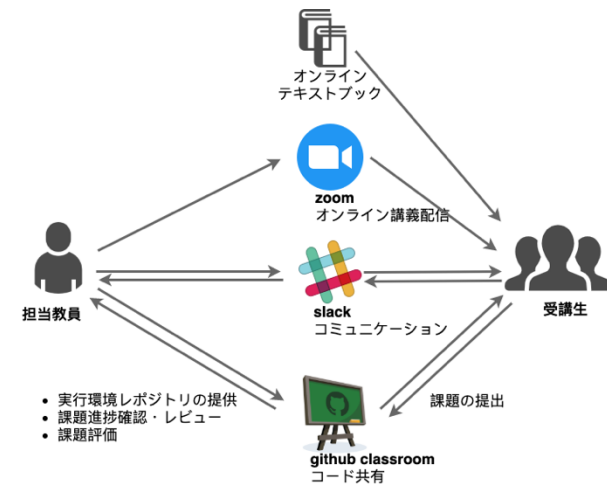


図 1. 講義の実施環境。

具体的なコンテナの構成は図 2 のとおりである。rocker/geospatial (Boettiger et al., 2020) は R と Rstudio server のみならず、地理空間分析に不可欠な GDAL (GDAL/OGR contributors, 2021) や PROJ (PROJ contributors, 2021) といった地理空間

データの入出力や座標系の変換を司るアプリケーションの他、tidyverse (Wickham et al., 2019) というデータ処理・分析のためのパッケージ群や、sf (Pebesma, 2018), tmap (Tennekes, 2018) といった R 上で地理空間データに特化した処理・分析・地図化を実施する各パッケージで構成されている。「2021GISClass_docker_renv」はこの rocker/geospatial をベースとし、Geocomputation with R のレポジトリ、課題提出用の関連ファイルをまとめている。

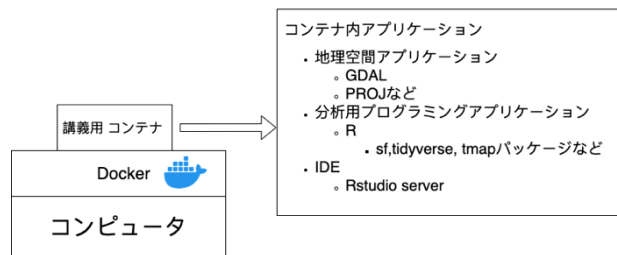


図 2. 講義で用いた Docker コンテナ概要。

4. 実習

R 入門では An Introduction to R から文法、データタイプの定義、データハンドリングといった、基礎的な内容をピックアップして授業で取り上げ、ハンズオン形式で講義中に教員がコードを書き実行させてみた。受講生は Docker コンテナを起動すると Rstudio server が利用することができるので、その環境でコードを書かせ実行させた。ジオコンピューテーション演習では、空間データハンドリング、座標系、空間演算、地図化を扱った。可能なところは講義中にコードを書き、実行例を提示したが、分析が複雑になるにつれ、コード作成やエラー処理に時間が取られることを想定し、レポジトリ内に配置した再現可能なコードを実行してみせ、解説を加える形とした。受講生も同様にコードを実行することができたが、自分でコードを書くプロセスを講義内で体験できなかった。そのため、講義後半に自由課題を設定した。自由課題では、講義で取り上げた内容を踏まえ、各受講生の興味のもとで対象を決定し、オープンデータをもちいて分析をし、結果を地図化した。

5. 成果物

html 形式のレポートとしたため、静的な地図のみならず、動的な地図を提出した受講生もいた。トピックは幅広く、感染症 (予測モデリング, 影響分析), 商圈分析 (コンビニエンスストアの出店戦略分析), 不動産 (物件立地評価), 交通 (カーシェアリング・シェアサイクルステーションマッピング), 災害 (ハザードマップ, 安全な避難分析), 地域社会 (子育てマップ), 環境 (気候変動影響分析), 犯罪 (自転車盗マッピング), 空間分析 (地理的加重回帰モデリング) などであった。

成果物の一例として子育てマップを紹介する (図3)。栃木県宇都宮市が提供しているオープンデータを用い、「保育園」、「幼稚園」、「赤ちゃんの駅」、「特色のある公園」を地図化した。R の mapview パッケージ

(<https://CRAN.R-project.org/package=mapview>) を活用することで動的な地図とし、地図の拡大・縮小を容易に実現している。地図化したことで市内にまんべんなく子育て関連施設が配置されていることがわかる一方、都心部に比べ郊外では車を所有する世帯も多く、利用者の特性にあわせた配置も考えていく必要があると考察している。なお、宇都宮市はまちかど情報マップの中でも子育てマップ

(https://www.machi-info.jp/machikado/utsunomiya_city/index.jsp?mode=2) という webGIS を公開しており、公的施設や観光施設などの情報を地図上で確認することができる。しかし、提出物の子育てマップはソースコードが約 20 行程度のものであり、プログラミング技術としても比較的容易なこと、必要であればデータの追加やレジェンドの変更など、カスタマイズ性が高いという利点がある。また html 形式となっているため容易にインターネット上に公開も可能である。

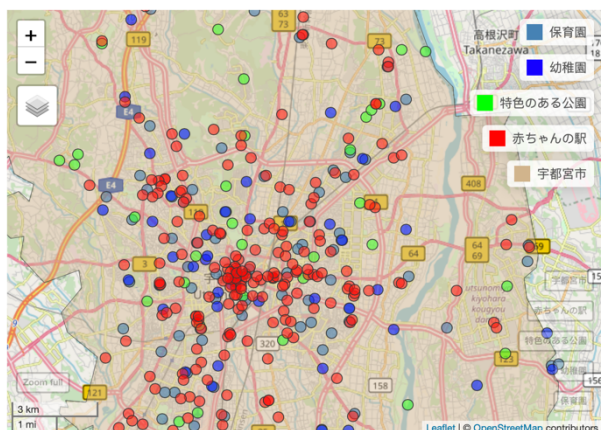


図3. 受講生により提出された最終課題の一例 (宇都宮市子育てマップ)。

6. おわりに

本稿では、オンラインで実施された地理情報科学特論の授業例を紹介した。Docker コンテナを活用することで、再現可能性の極めて高い R 実行環境を受講生と共有することが可能となった。また、Github Classroom でのソースコードの配布と課題の提出を実施したことにより、Github をベースとした課題管理・進捗管理が可能であった。大学が提供している学習管理システムでは成果物の提出管理ができるものの、途中経過を確認することが難しい。Github Classroom では、受講生が取り組み中の課題であっても commit, push が実行されれば教員に通知されるため、どの程度の進捗であるか、どこでつまづいているかが容易に確認することができる。これら一連のオンライン環境はスケーラブルであり、基本的には受講生の数に関わらず対応可能である。ただし、学生へのきめ細やかな対応を実現するには、これ以上の受講生の増加にはティーチングアシスタントによるサポートが必要となるだろう。

再現可能性という点では利点のある仕組みと考える反面、GIS 教育としては課題も残る。GUI の GIS と比べ、分析の一つ一つのプロセスを地図化する必要がないため、プログラムの中でどのような処理を実施しているのか、理解が難しい場合もあるだろう。分析の都度、結果を地図化すればよいのだろうが、CLI だと地図化も一手間かかるため、初学者には難しい場合も多いかもしれない。この点については、近年 R Shiny

(<https://CRAN.R-project.org/package=shiny>) の枠組みにおいてインタラクティブな GIS ツールの構築が可能となってきており、さらなる技術的発展が望まれる。

また、今回は R を用いた授業となったが、受講生のほとんどが R に触れたことがないため、R の基礎から実施する必要があった。よりユーザー数の多い python を用いた同様の講義の実現も考慮する必要があるだろう。

本授業は情報領域での開講であったため、受講生の多くはプログラミングや情報技術に抵抗が少なく、特段のトラブルもなく進めることができた。今後も地理情報を活用した各受講生のアイデアを具

現代できるアプローチを講義内で提供していきたい。

注釈

*1 Github の一操作であり, Github 上にある自分以外のレポジトリを自身の Github レポジトリとしてコピーする際に用いられる。

*2 Git コマンドの一つであり, 対象ディレクトリ内で変更, 追加, 削除されたファイルを git で管理する際に用いられる。どのような変更が行われたかを把握するためのコメントを付加することができる。

*3 Git コマンドの一つであり, 作業を行っているコンピュータ環境内 (ローカル) の対象ディレクトリの変更履歴を Github をはじめとするリモートレポジトリにアップロードする際に用いられる。

これによりローカルでの作業をクラウド上で管理することができる。

*4 Dockerfile をもとに構築されたレイヤー群であり, コンテナを構築するために用いられる。

*5 仮想化技術の一種であり, Docker engine のようなコンテナ管理アプリケーション上で動くことのできるアプリケーションである。Virtual machine のような仮想化技術に比べ, OS に含まれる不必要なアプリケーションを排除することができるため, 小規模化が実現できる。また, コンテナ管理アプリケーションがあればどの環境においても同一コンテナを起動することができる利点を有する。

*6 Dockerfile にアプリケーションのインストールコマンドなどが記載されている場合, コンテナ構築のタイミングにより, 完全に一致したバージョンのアプリケーション群が入手できない可能性がある。そのような問題は, 前もって作成した Docker image を配布する事により回避することができる。

*7 2021年4月時点では Rstudio server が M1 チップが搭載された macOS 上での Docker では起動しない問題があったため, 別レポジトリ

「2021GISciClass_docker_renv_m1」 (Tsutsumida, 2021b)を用意し対応した。

謝辞

Geocomputation with R の著者である Robin Lovelace, Jakub Nowosad, Jannes Muenchow, An Introduction to R の著者である Alex Douglas, Deon Roos, Francesca Mancini, Ana Couto, David Lusseau には著書をオーブ

ンライセンスでオンライン公開していただいたため本講義の教科書として採用することができた。心より感謝の意を表す。

また, 本講義に参加してくれた受講生には, このような新たな取組にも関わらず授業に参加・貢献していただいた。少しでも皆さんにとって有意義な授業となっていたことを期待する。

参考文献

- Abrahart, R.J., Openshaw, S., See, L.M. (2000) *GeoComputation*, CRC Press.
- Boettiger, C., Robin, Howe, M., Lamb, J. (2020) Rocker-org/geospatial: 3.6.2., Zenodo, <https://doi.org/10.5281/zenodo.3736671>
- Brunsdon, C., Singleton, A. (2015) *Geocomputation: A practical primer*, Sage.
- Douglas, A., Roos, D., Mancini, F., Couto, A., Lusseau, D. (2021) An introduction to R, <https://intro2r.com>.
- GDAL/OGR contributors (2021) GDAL/OGR geospatial data abstraction software library, Open Source Geospatial Foundation.
- Longley, P.A., Brooks, S., Macmillan, W., McDonnell, R. (1998) *Geocomputation: A primer*, Wiley.
- Lovelace, R., Nowosad, J., Muenchow, J. (2019) *Geocomputation with R*, CRC Press.
- Pebesma, E. (2018) Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* 10, 439–446.
- Peng, R.D. (2011) Reproducible research in computational science, *Science* 334, 1226–1227.
- PROJ contributors (2021) PROJ coordinate transformation software library, Open Source Geospatial Foundation.
- QGIS Development Team (2009) QGIS geographic information system, Open Source Geospatial Foundation.
- R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Environmental Systems Research Institute (2011) ArcGIS desktop: Release 10, Redlands, CA.
- Tennekes, M. (2018) tmap: Thematic maps in R. *Journal of Statistical Software* 84, 1–39.
- Tsutsumida, N. (2021a) 2021GISciClass_docker_renv, Zenodo, <https://doi.org/10.5281/zenodo.5642590>

ポストプリント版

堤田成政 (2022) 再現可能性を重視したオンラインでのジオコンピューテーション教育, GIS -理論と応用 vol. 29(2) pp. 59-64.

Tsutsumida, N. (2021b)

2021GISciClass_docker_renv_m1, Zenodo,

<https://doi.org/10.5281/zenodo.5642528>

Wickham, H., Averick, M., Bryan, J., Chang, W.,
McGowan, L.D., François, R., Golemund, G., Hayes,
A., Henry, L., Hester, J., Kuhn, M., Pedersen, T.L.,
Miller, E., Bache, S.M., Müller, K., Ooms, J.,
Robinson, D., Seidel, D.P., Spinu, V., Takahashi, K.,
Vaughan, D., Wilke, C., Woo, K., Yutani, H. (2019)

Welcome to the tidyverse, *Journal of Open Source
Software* 4, 1686.

Xie, Y. (2016) *Bookdown: Authoring books and
technical documents with Rmarkdown*, Chapman;
Hall/CRC, Boca Raton, Florida.

Xie, Y., Dervieux, C., Riederer, E. (2020) *R markdown
cookbook*, Chapman; Hall/CRC, Boca Raton, Florida.