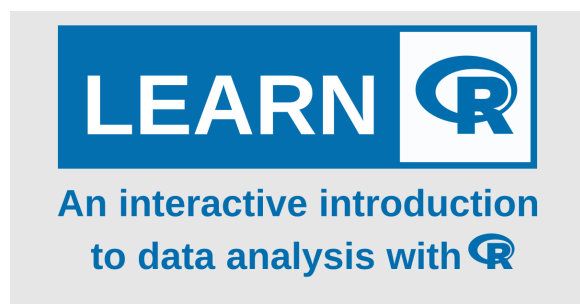


Written Statement for the Dean's Award for Innovation in Learning and Teaching

BeginR dev team (Danyang Dai, Jessica Leung, Mitchell O'Hara-Wild, Emi Tanaka)



Synopsis

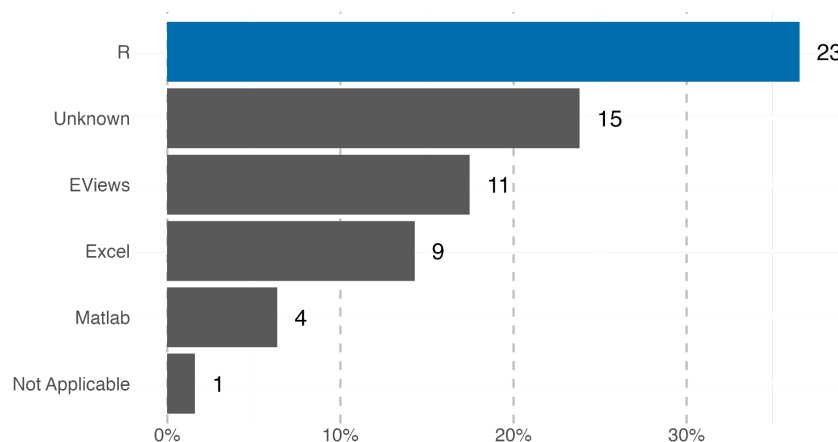
Programming is a core skill to flexibly manipulate, visualise and fit models to various forms of data. R is one of the top programming languages used by analysts and alike, and is the primary programming language used across the units offered by the Department of Econometrics and Business Statistics (EBS). Students may not have sufficient skills needed to do data analysis with R; this poses a dire problem for both instructors and students in units that cannot dedicate the resources to teach R, which are often peripheral to the unit's core aim, yet necessary for computing sections of the unit. We developed the "Learn R" website (<https://learnr.numbat.space/>) as a supplementary, interactive, self-paced learning tool for students to refresh or learn R for major data analysis tasks. This innovative resource provides a cost-effective solution to improve student learning experience across multiple units offered by EBS.

Background

Most EBS units require students to demonstrate statistical thinking through application. The application is based on a particular software, which in EBS units varies from a spreadsheets tool, Excel, a graphical user interface tool, EViews, a proprietary tool, Matlab and a programming language, R. Our internal survey in 2021 (with results shown in the graph below) revealed that at least 23 EBS units use R with some considering replacing EViews to R; the second most used software was EViews (at least 11 units) followed by Excel (at least 9 units) and then Matlab (at least 4 units). This makes R the predominant software in EBS units and with the trending increase of usage in R for our units, there was a pressing need to bring higher year students without sufficient R knowledge up-to-speed with R and still deliver the main content of the unit.

Software usage across EBS units from 2021 internal survey

Most EBS units use R



Note: some units use multiple software

To fill the gap of R knowledge in higher year units, Dr. Emi Tanaka was originally tasked to run a one-off workshop just prior to the semester, however, she did not see this as a sufficient solution for continuous learning nor a method that scales well with the number of students. Indeed, a number of literature, such as Bashir & Hoque

(2016) and Robinson & Carroll (2017), propose the use of e-learning or self-paced online learning specifically for programming and outline the benefits for engaging students for self-learning and life-long learning, whilst also significantly reducing the burden of work on the unit instructors.

Our Contributions

As an alternative proposal, Dr. Tanaka assembled a team (initially Danyang Dai and Mitchell O'Hara-Wild with support of departmental funds, and later Jessical Leung who joined as a new faculty member in 2022) to develop ideas and resources for learning R. As well documented in educational literature, such as Tullis & Benjamin (2011) which outline the effectiveness of self-paced learning, we quickly settled on creating a self-paced online website with quizzes that assess knowledge for instant feedback. As suggested by the National Research Council (2000), students are motivated when they see the value and utility of what they are learning, so we widely consulted the EBS staff on the scope of the topic and format to be delivered so the developed content will be relevant for students. The circulated planning document, which shows the topics, tasks and schedule, can be found in the link in the Supporting Documents with a screenshot shown in **Supporting Documents Figure 1**.

The result of the consultation and development resulted in 8 topics:

- | | |
|---|------------------------------|
| 1. Getting started with R & RStudio | 5. Data wrangling with R |
| 2. R programming basics | 6. Data visualisation with R |
| 3. Troubleshooting and asking for help | 7. Linear regression with R |
| 4. Project structure and importing data | 8. R Markdown basics |

hosted at <https://learnr.numbat.space/> (which we will henceforth refer to as the **Learn R website**) with a screenshot shown in the **Supporting Documents Figure 2**. This development was a huge team effort; Mr. O'Hara-Wild led the development of the website framework and each of the other team members contributed content for 2-3 topics. All materials to generate the website is based on open-source tools hosted on GitHub at <https://github.com/numbats/learnr>. This open-source format adds transparency in the process and makes it easier to invite others to contribute – both of which were important considerations for future maintenance of this resource.

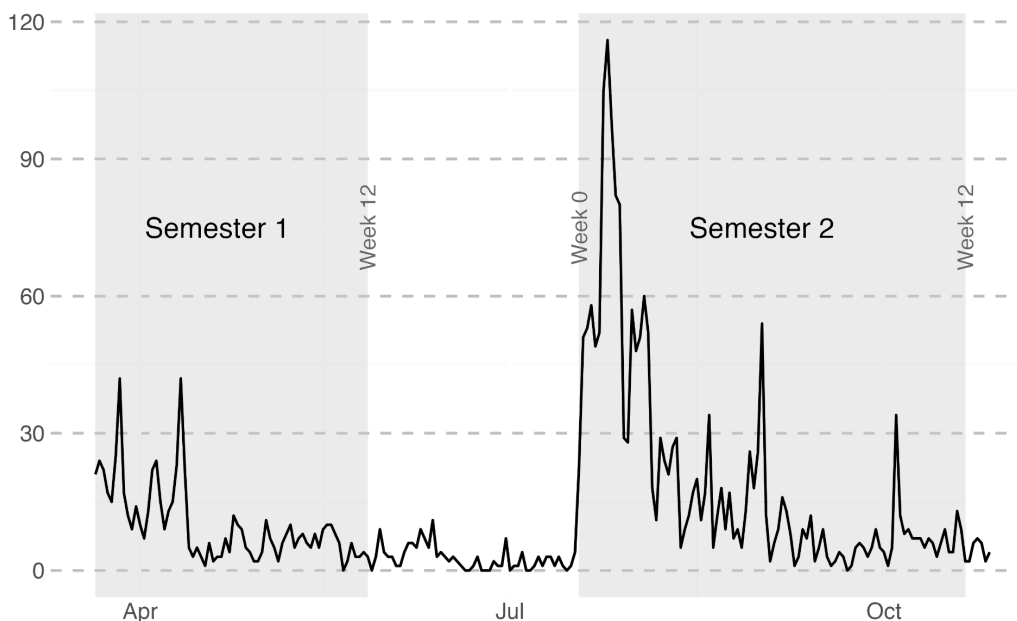
The content was designed for self-directed learning that blended explanations, examples and quizzes as shown in **Supporting Documents Figures 3 and 4**. More specifically, the platform consists of a number of modules which first introduces certain topics in R and then prompts students to perform coding tasks to move on. Students receive instant feedback for coding exercises – when there are errors in their code, an authentic R error message will be provided. Students will be guided with hints and feedback to arrive at the correct way of coding the required task. The provision of this support puts students back to the centre of their learning experience by giving them the opportunity to choose what to learn and when to learn. The self-paced element also encourages independence in learning and fostering the growth mindset. It is also an authentic learning experience since modern programming language skills are often acquired by performing coding tasks (Bashir & Hoque, 2016 and Robinson & Carroll, 2017).

The information about the Learn R website was distributed to EBS staff members. Those staff members who opted-in to use the website directed the students to the relevant topics within the Learn R website. This resource is now used as supplementary resource in many of the EBS units, including ETC1010/5510 Introduction to Data Analysis, ETF2480 Business Modelling, ETC2420/5242 Statistical Thinking, ETC3550 Applied Forecasting, ETC5523 Communicating with Data, and so on. As attested by Ms Dai, who tutored for ETC2420/5242 before and after the usage of the Learn R website, there was a noticeable reduction in the need to help students in basic installation of R (covered in Chapter 1 of the Learn R website). From the number of testimonies we have heard, we believe the Learn R website has improved both the teaching experiences of staff and learning experiences for students.

Evidences

We started collecting the Google Analytics data of the Learn R website from the 21st March 2022. The **Supporting Documents Google Analytics Data** shows the link to the interactive data report as well as the screenshot of the report from the Google Analytics platform that shows there were approximately **2,000 users that engage on the site for about 8 minutes** on average. The figure below shows the number of users that visited the Learn R website from 21st March 2022 to 27th October 2022 with annotations for the semester period. We clearly see that the number of users coincides with the semester period with the decline of users expected in later weeks as students master the basics of R.

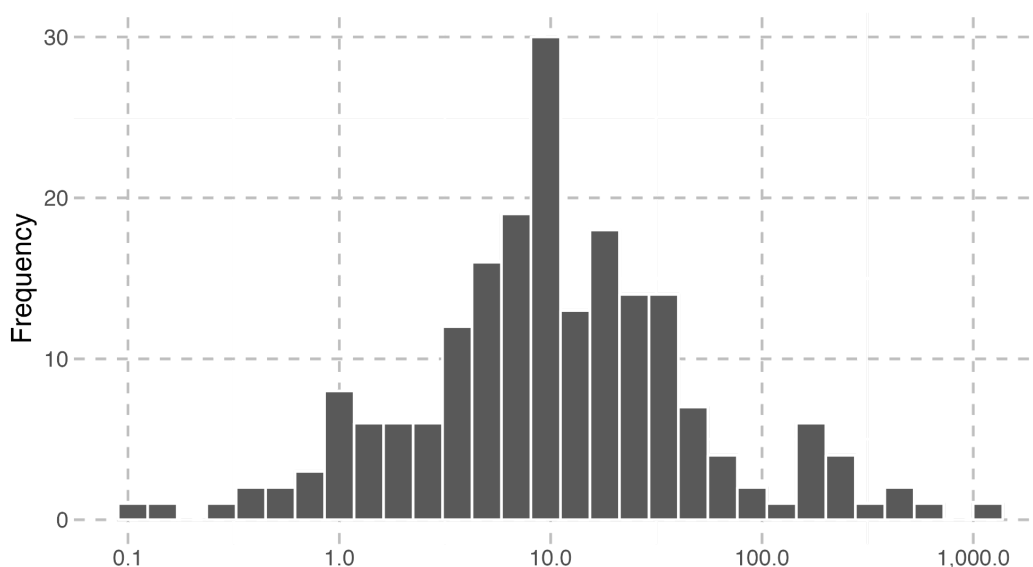
Number of users for the Learn R website



The figure below shows the histogram of the average engagement time per user across the period of 21st March 2022 to 27th October 2022. We can see from this graph that the mode of average engagement time per user is about 8-10 minutes, but there are as many users that spend well above 10 minutes on the site.

Average engagement time per user across days (in minutes)

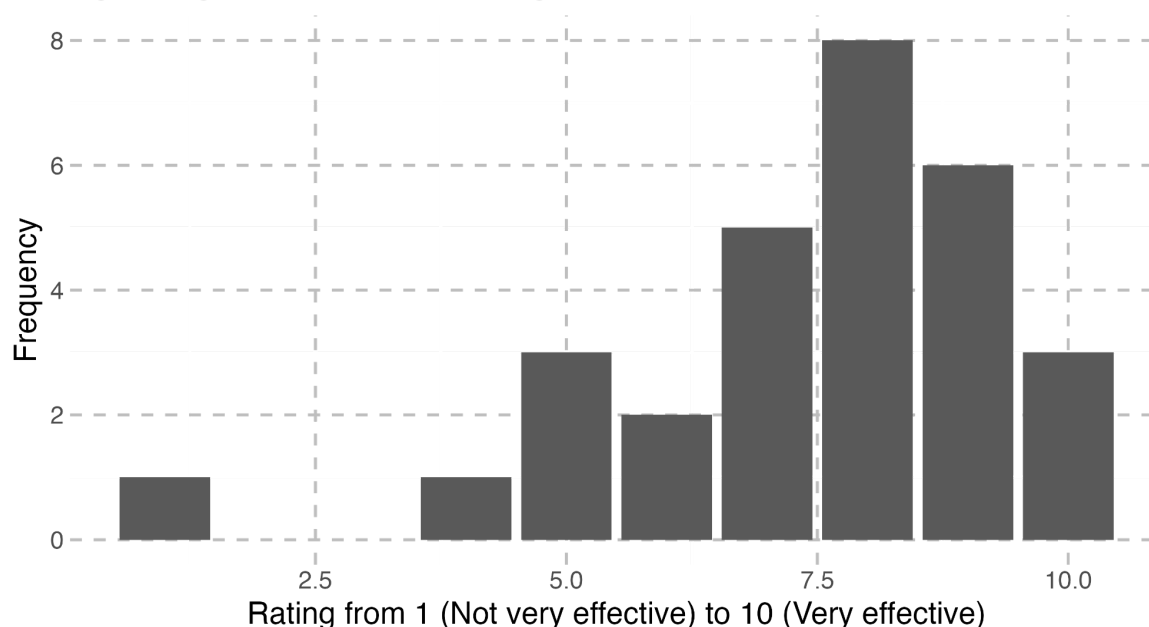
Most users spend about 8 minutes on the website



We also conducted a survey of the students that ran from 29th September 2022 by asking Chief Examiner's to distribute the Google Form survey found [here](https://forms.gle/EVNFmdHVS6vwFeUF8) (<https://forms.gle/EVNFmdHVS6vwFeUF8>) to their students. The full data from the survey is available in the link provided in **Supporting Documents Student Survey Data**. We unfortunately did not receive many responses (32 in total). The graph below shows the frequency plot of the student's rating on overall effectiveness of the Learn R website. Noticeably, we had one observation with a rating of 1 (Not very effective) but upon inspection of the comments, we realise that the student is not commenting on the Learn R website but about their unit. The qualitative comment suggested some improvements we can make (e.g. "modules are too long") but the vast majority did not have an issue, was content with it ("The LearnR website itself has been very concise and clear"), or found it helpful ("This was really cool! I really appreciated this resource, I'd already done some R beforehand, but I really learnt a lot and had some great refreshers and learnt some more best practices from it."). These qualitative feedback matched the quantitative rating, where average is 7.4 and median is 8.0 out of 10. This survey demonstrated to us that the Learn R website has been quite effective with some room for improvements.

How would you rate the overall effectiveness of this resource for you?

Average rating is 7.4 and median rating is 8.0 out of 10.0



References

- Bashir, G.M.M. & Hoque, A.S.M.L. (2016) An effective learning and teaching model for programming languages. *J. Comput. Educ.* 3, 413–437.
- National Research Council. (2000). *How People Learn: Brain, Mind, Experience, and School: Expanded Edition*. Washington, DC: National Academy Press.
- Robinson, P. E. and Carroll, J. (2017) An online learning platform for teaching, learning, and assessment of programming. *IEEE Global Engineering Education Conference (EDUCON)* 547-556
- Tullis, J. G. and Benjamin, A. S. (2011) On the effectiveness of self-paced learning. *J Mem Lang* 64(2) 109-118