Improving Statistics Education Through Interactive Learning Tools

Philipp Burckhardt* Alexandra Chouldechova*

Abstract

We introduce an e-learning platform called ISLE (*Interactive Statistics Learning Environment*) that provides a framework for building interactive online lessons for statistics that can be used in a blended-learning setting. The platform comes with an accompanying analytics dashboard that enables instructors to easily track the learning trajectories of their students. In this paper, we reflect on an analysis of student engagement with interactive lab sessions that were administered over three sections of a half-semester course on R programming and data analytics. Having collected almost $25,000$ user interactions, we analyze the completion rates of the labs via group-based trajectory modeling. We identify distinct student groups who follow a similar path over time, and investigate the correspondence between learning trajectories and learning outcomes. Taking a closer look at click rates within the R exercises, we find that not all students approach the problems of a lab session in a linear manner. We discuss the implications of these preliminary results and close by laying out future directions for the ISLE platform.

Key Words: Education, e-learning, machine learning, teaching, statistics, trajectories

1. Introduction

Student participation in hands-on laboratory activities has a long tradition in the physical and natural sciences. While the science education literature has produced mixed results on the benefits of such hands-on activities [1], there is evidence to suggest that exercises which allow students to manipulate experimental factors can improve learning outcomes [2].

Despite the similarities that statistics bears to other sciences, lab-style activities are still rather uncommon in statistics classrooms. The typical statistics class tends to bear a closer resemblance to courses in mathematics, despite the oft-heard protestation that statistics is not math. Indeed, at the introductory level, statistics arguably has more in common with computer science, a field where it has been suggested that when it comes to labs, the more the merrier [3, 4]. The success of lab activities these closely related fields motivated our work on facilitating similar learning experience for students in the data sciences.

We developed the ISLE (*Interactive Statistics Learning Environment*) framework to allow instructors to create and deploy hands-on e-learning modules for classes in data analytics, statistics, and statistical computing. This system offers several key advantages over more traditional approaches to statistics education.

Content creation and sharing: Since the created e-learning modules can be fed into a public gallery that is accessible through an online dashboard, they are available for any instructor, who may then use them in his or her own courses, either in an unmodified or modified form. This intersubjectivity fosters collaboration and pedagogical best practices to emerge and gain traction.

What does a good lab look like?: What makes for a good statistics lab remains an open question. Research suggests that classical hands-on activities such as coin flipping and die rolling fail to improve learning outcomes [5]. This may seem somewhat at odds with the ample evidence that interactive activities can have a much higher impact on student learning than video lectures or reading materials alone [6]. We argue that activities such as the ones mentioned above are too far removed from statistical practice.

*Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213

ISLE instead takes advantage of visualizations, simulations, and real-world data sets, prompting students to engage with data in a goal-oriented way. ISLE activities can be made to more closely mimic guided real world statistical and data analytic tasks, and are therefore much more closely connected to statistical practice. Drawing upon the research on Intelligent Tutoring Systems, which have been shown to have a positive effect on students understanding and retention of subject material [7], ISLE provides facilities for personalized feedback. Moreover, we can use the user activity logs to learn whether students are interacting with the modules in the manner intended, and to run A/B tests to identify module changes that help to improve engagement.

What does success look like?: Our framework allows the logging of various features of how students interact with the modules (e.g., mouse clicks, scrolling, durations spent on an item). As we show in Section 5, analyzing these logs provides insight into common patterns of engagement, along with a sense of the kinds of metrics that are associated with good learning outcomes. What we learn about meaningful engagement metrics feeds directly into the design of the ISLE analytics dashboard, which allows instructors to easily perform similar analyses of the ISLE data from their classes.

1.1 Outline

The remainder of the paper is structured as follows: In the next section, we provide some background on online learning and blended learning platforms. Then in section 3 we describe the ISLE project, an open-source learning platform, developed and actively used at Carnegie Mellon University. In Section 4, we describe the data that we have collected through ISLE-powered lessons. The collected interactions span structured data (e.g., time stamped clicks, inputs, etc.) as well as unstructured data (such as free text and code). In the following methods section, we discuss different measures of student engagement that can be used to create a user model and discern personalized learning paths. As an underlying machine learning algorithm, we rely on group-based trajectory modeling. In the results section, we present preliminary insights that we have gained into learning behavior and which may influence student performance. In Section 8, we discuss our findings, their limitations, and future perspectives.

2. Blended learning in Context

Historically, the advent of a new medium that could be utilized for educational purposes usually undergoes several reaction phases. At first, there is skepticism due to the novelty. This is often followed by an inflated sense of what learning materials from the new medium can accomplish, which is tempered when the community gains an understanding of what can and cannot be achieved in said medium. A recent large meta-analysis of online learning has demonstrated its advantages compared to face-to-face instruction, but the authors caution that this might be more due to blended learning than the online medium alone [8].

Similarly, Clark and Meyer reflect on the research surrounding the effectiveness of online learning, and demonstrate that learning is achieved through adherence to evidence-based instructional design principles and best practices, not simply by converting the material from, say, a textbook into a video course [9]. While we share this perspective, it is important to highlight that online learning platforms provide one key advantage that sets them apart from previous approaches, namely the fact that one can monitor, store, and analyze student interactions, which can – through a process of iterative design – be used to improve instruction. This has become visible in the rise of MOOCs (massive open online course) over the last few years, which has sparked new research on how to make sense of the vast amounts of collected data. However, in an academic environment, online learning materials have to blend with traditional, face-to-face modes of instruction.

To differentiate this new blended learning approach that combines the advantages of the virtual with the physical classroom, Armando Fox, directory of the Berkeley MOOC lab, has coined the term SPOC (small private online course) [10]. The Open Educational Resources (OER) movement, which sparked projects such as MERLOT from the California State University, OpenCourseWare (OCW) by the Massachusetts Institute of Technology (MIT), the Carnegie Mellon Open Learning Initiative, among many others, has emphasized the free redistribution, remixing, and reuse of learning material. Since this creative commons philosophy has found its technological counterpart in the open-source software movement, rapid development of high-quality interactive e-learning content that can be used for SPOCs is increasingly feasible.

3. ISLE

In this section, we describe ISLE (Interactive Statistics Learning Environment) from an infrastructure and learning analytics perspective. At its core, ISLE is a framework which – by means of its accompanying editor – allows for the easy construction of statistics-related e-learning modules. ISLE is open-source (MIT licensed) and can be downloaded at the following address: `http://isledocs.com`.

3.1 Research Laboratory

One goal of the ISLE project is to build up a research laboratory that allows instructors to better monitor their classes and understand how students interact with class material. As an instructor, one is commonly restricted to a small set of observable learning assessment events (exams, quizzes etc.). Furthermore, instructors usually do not grade homework, but delegate this task to teaching assistants. Often, the only work that an instructor will see from a student is his or her final exam. In sharp contrast to this, all user interactions are monitored when students interact with ISLE lessons and can be analyzed through an accompanying dashboard. This turns the classroom into a research laboratory. We hypothesize that a system of feedback collection integrated into the instruction and monitoring of user actions can reveal issues students are facing and uncover what material they find interesting or uninteresting. To test this hypothesis, we are gathering data on the time spent in the lessons, all mouse clicks, inputs to text fields, interactions, and feedback provided through an integrated feedback submission system.

3.2 Features of ISLE

Built using advanced browser technologies, ISLE lessons run natively in a web browser without the need of any plug-ins. Using responsive design, ISLE lessons do not only work on all operating systems, but also on handheld devices such as iPads or mobile phones. Since each ISLE lesson is a single file, no complex directory structures need to be managed. As ISLE is built using web technologies, it is possible to use JavaScript, HTML and CSS for full customizability and a platform-independent codebase. Using the facebook React.js library for the user interface, ISLE lessons can be built by combining reusable and configurable building blocks comparable to Lego bricks. Built on top of the open-source numerical computing library stdlib for JavaScript, many calculations can be done on the client-side, allowing for interactive visualizations and simulations without any server communication. The OpenCPU system written by Jeroen Ooms at UCLA [11] provides a gateway to include code written in the statistical programming language R. Since e-learning allows recording, storage and replay of user interaction, it is possible to conduct a fine-grained assessment of student learning and the built lessons. The ISLE dashboard gives instructors the capabilities of hosting ISLE lessons (comparable to `shinyapps.io` for interactive R widgets) plus easy access to students personal learning paths and progress statistics.

3.3 Iterative Design

The use of ISLE lends itself to a process of iterative design, where usability, adaptivity, and scalability are continually assessed and improved. In its first year, ISLE has been used in classes for Masters students enrolled in the Heinz College, the Public Policy and Information Systems School at Carnegie Mellon University. Starting this Fall, ISLE will be incorporated into a new undergraduate introductory data science class at the Department of Statistics that is aimed at a broad audience spanning a wide range of majors. Furthermore, we are planning to develop ISLE lessons for professionals from the healthcare sector, who are undergoing on-the-job training. Combining online learning aspects with vivid classroom instruction in a blended learning context allows that the two spheres (online/offline) may complement each other instead of being treated as mutually exclusive.

4. Data

In class 94-842: Programming in R for Analytics taught by A.C. at the Heinz College, we deployed ISLE lessons in the first and second mini (half-semester) of the Fall 2016 semester as well as the first mini of Spring 2017. Students were exposed to twelve interactive labs, a pre-test assessing their statistics knowledge going into the analytics component of the class, and three follow-up multiple-choice quizzes, which tested the conceptual understanding of the covered topics. The schedule for the course, which spans eight weeks, is displayed in the Figure below[1]. Most of the students taking the course were enrolled in the Masters programs for either information systems or public policy at the Heinz College.

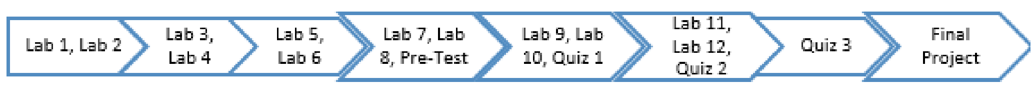

Figure 1: Schedule for the course "94-842: Programming in R for Analytics". Each box represents a single week of the course. The boxes contain all ISLE activities occurring in the respective week. In week eight, there were no activities involving ISLE-powered lessons, as students were doing their final projects.

The labs consist of a mix of interactive, explorable simulations explaining the underlying statistical concepts, R exercises with accompanying hints to practice programming, and multiple-choice and free-text questions for students to test their understanding. The mean number of questions posed to students per lab is 13.4 (excluding quizzes). Participation credit was given to the students for attempting the labs, but completion of them was completely voluntary.

In the first iteration of the course (mini 1), labs were taken during the last twenty minutes of class, with students usually staying longer to finish them in case the amount of time during class was insufficient. Teaching assistants were present to assist with any open

[1]The course webpage, which includes a detailed description of the materials covered in each week, is available at the following address: `http://www.andrew.cmu.edu/user/achoulde/94842/`

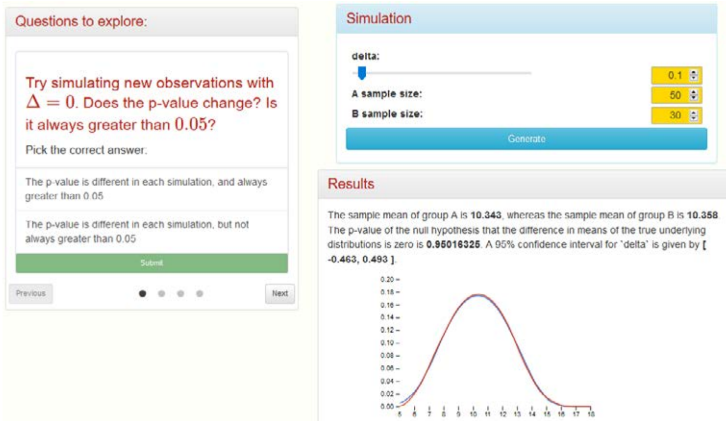

Figure 2: Simulation study for a comparison of means from Lab 7.

questions. In contrast, students took the labs at their own time and pace during the second and third iteration of the course. On average, we have data for seven labs per student.

4.1 Student Actions

Over the entire time frame, we have collected $24,500$ student interactions with the platform, belonging to 172 unique students enrolled. We define student interactions as all clicks, inputs to text fields, answers to R exercises etc. Each action is associated with a time stamp. Overall, the mean number of actions performed by students in a single session was 22.2.

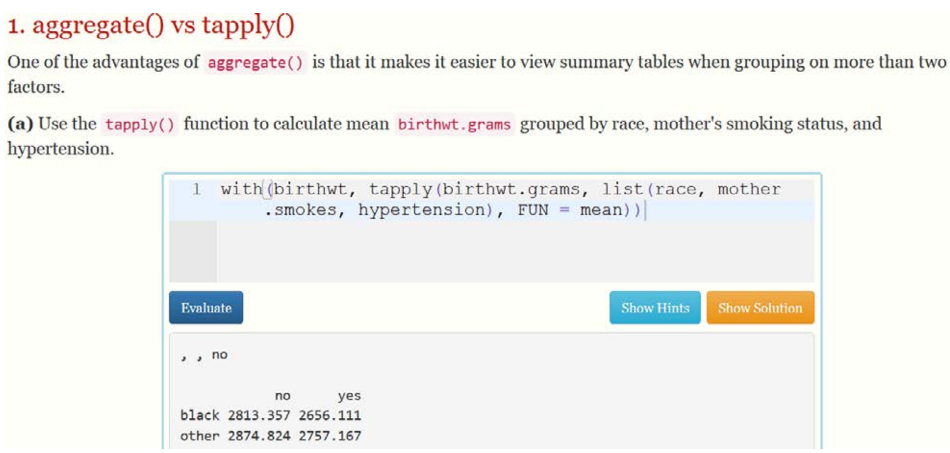

Figure 3: Screenshot of an R exercise from Lab 3. The R shell allows students to execute R commands and comes with accompanying hints and a button to compare the submitted answer with an accepted solution.

4.2 Student Assessments

The courses concluded with a final project that spanned the whole curriculum and had the students analyze a data set and write up a report. Out of a total of 45 points, students obtained an average grade of 36.3 points. Besides this summative assessment, we have the student answers to the multiple-choice questions on the quizzes. Each quiz (besides the pre-test) contained a total of eight exercises. Besides these two instances of formative and summative assessment, students also had to complete weekly homework assignments.

5. Methods

Group-based trajectory models (GBTMs) allow to identify groups of individuals who share a common trajectory of some outcome over time. Originally devised by Nagin and colleagues as a quantitative method to identify different trajectories in criminal careers, they can be used as well to track the behavior and progress of students in an educational setting [12]. At their core, GBTMs are semi-parametric mixture models, in which for each one of a finite number of latent groups a response variable is modeled as a function of time and potentially other covariates.

5.1 Measures of Engagement

There is a variety of measures that one could track to monitor how students engage with the online labs. However, all these are only approximations to what we really would like to measure, the active engagement with the labs and the contained exercises.

5.1.1 Time Spent

One such outcome measure is the amount of time spent by the students inside the labs. However, we found that many students finished the labs only after long time periods without any activity, presumably because they either read up about the covered topics or got distracted and browsed to unrelated websites. While there are remedies to deal with this ex post such as trimming the data to exclude unrealistic values, most of these solutions are rooted in subjective judgments and might be called into question on that basis. Going forward, we will advance the platform by only measuring the time the students have the lab windows in focus inside of their browsers, and not while they spend their time elsewhere.

5.1.2 Number of Actions

Another measure one could consider is the number of actions students perform in each lab. However, this is an imperfect metric likewise as more clicks and form submissions dont necessarily map to better student engagement with the subject material. A difference in the number of actions could simply be due to different styles of using our platform: Some students might think longer about a problem before attempting it, while others will evaluate several chunks of R code in rapid succession while coming up with their answer. Looking at the data, one can confirm that actions are often clustered with each other, which lends support to this hypothesis. We present some findings in this direction in Section 5.2.

5.1.3 Completion Rates

For the present analysis, we have decided to focus on the proportion of exercises inside the labs that were attempted by the students since it does not suffer from the shortcomings as the two previously discussed outcome variables. On the other hand, should all students complete all exercises, this might not give us very fine-grained insights. The actual numbers show that students on average solved roughly 82% of all exercises, with the median lying above 90%. Hence going forward, we will try to investigate how to refine the measurement of student engagement.

5.2 Model Formulation

Let $\boldsymbol{y} = (y_1, \ldots, y_T)^\intercal$ be the vector of T observations of an outcome, such as the lab completion rate. In the mixture model formulation of a GBTM, the conditional density h of outcome trajectory y is modeled as a weighted sum of K components, i.e.

$$h(y \mid x, \pi, \theta) = \sum_{k=1}^{K} \pi_k f(y \mid x, \theta_k), \tag{1}$$

where the class probabilities satisfy $\pi_k > 0$ and $\sum_{k=1}^{K} \pi_k = 1$. The main simplifying assumption of GBTMs is that conditional on group membership, observations at different points in time are independent of each other. In mathematical terms, this means that the joint density inside each group factorizes as follows over the T available time periods:

$$f(\boldsymbol{y} \mid x, \theta_k) = \prod_{t}^{T} f(y_t \mid x_t, \theta_k), \tag{2}$$

where y_t denotes the measurement of the outcome at time t, x_t the covariates, and θ_k the corresponding group-specific coefficients. When f is assumed to be a Gaussian density, this model is a standard mixture of linear regression models applied to longitudinal data. For other members of the exponential family, it becomes a mixture of generalized linear models [13]. In our case, the outcome y_t corresponds to the number of questions attempted by a student in each lab. It is sensible to treat the number of attemped questions as binomially distributed. With the logit as the chosen link function, this corresponds to modeling the probability p_t of attempting a question as

$$p_t = \frac{1}{1 + \exp\left(-g(x_t)\theta_k\right)}, \tag{3}$$

where we use a cubic spline for $g(x_t)$ to allow for some flexibility in the estimated trajectories, with coefficients for each combination of minis (I, II, or III) and the two grade groups. This model estimates a total of 24 coefficients in each of the two groups.

6. Results

Our preliminary results can be summarized as follows:

- The trajectory model allows us to identify two distinct groups of students who differ in their completion rates over time. While one group consists of students who complete most of the exercises, a minority of students do not. Section 7.1 elaborates on these findings and the shapes of the fitted trajectories.

- Looking at the click rates for individual R exercises in Section 7.2, we find that students complete exercises in a non-linear fashion. Not unexpectedly, harder questions take considerably more tries, but the data also indicates that students are much more prone to delay or revisit such exercises.

6.1 Group-based Trajectory Modeling for Identifying Patterns in Student Engagement

Due to the limited number of students present in our data set, we have fitted a model with just two latent groups, which should allow us to uncover in broad strokes the main archetypical trajectories. The results of this alongside the original data are displayed in Figure 4.

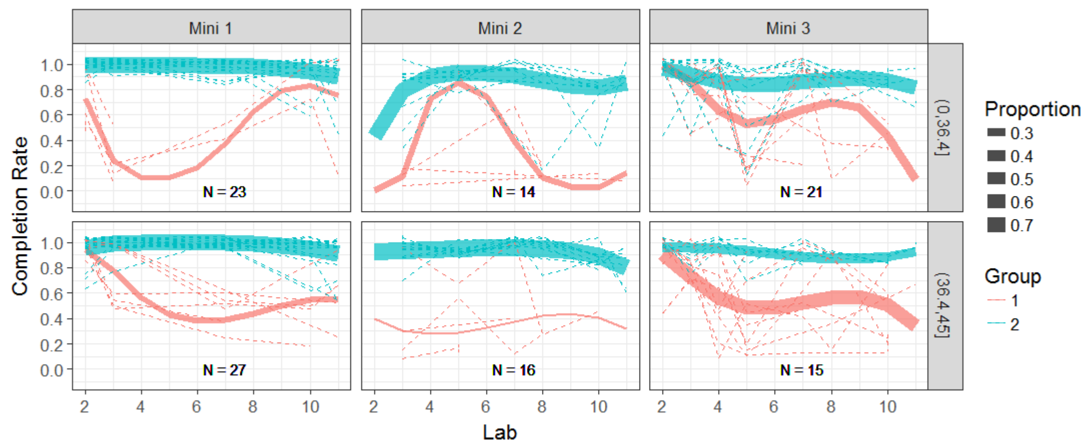

Figure 4: Trajectories for the completion rates of each lab by the students, split by class section (mini) and whether students performed below average (36.4) on the final project. The solid lines represent the fitted trajectories of a group-based trajectory model with two latent groups. Each trajectory is modeled as a cubic spline, with line thickness indicating the proportion of students from the respective categories who were assigned to the given group. Dotted lines show the observed individual student trajectories, which have been jittered to avoid over-plotting. In all plots, we see one group characterized by a completion rate close to one for most labs, and one other group with a trajectory that is first decreasing before it starts increasing again as the semester reaches its end (the plot in the middle of row one is the exception). The second group could be comprised of students who lose motivation early on before they spend more time again on the course as quizzes and the final project draw nearer. The interpretation of the trajectories is discussed in Section 5.1.

The solid lines showing the fitted trajectories reveal that there seem to be differences both across the different minis[2] as well as the final grades of the students. Students assigned to the blue group for the most part completed all the questions inside the labs, although there are some differences across the different sections of the course: Whereas the line appears to be straight for the first mini when students took the labs inside the class room, there is a slight trend in both the second and third minis. And indeed: the interaction terms of the mini and grade indicator variables with the trajectories are significant at the 1% level. In contrast to the students assigned to the blue group, the red line shows the fitted trajectory for those who completed only a smaller amount of the exercises in each lab. Rootograms[3] of the posterior class probabilities show that the student trajectories can be well separated between the two groups. As we can see from Figure 4, the trajectories for the red group are not constant, instead showing for most categories first a decrease over time followed by an increase and sometimes another

7. A Closer Look at Click Rates

One finding of our preliminary investigation into the data is that students approach the exercises of the lessons in a non-linear manner. The lab sessions for the R Programming class were designed as single-page applications unrolling as a sequence of individual exercises, frequently building on top of each other. Considering this, one might expect that students

[2]94-842 is a half-semester class. Half-semester terms at the authors institution are referred to as minis.

[3]Like a histogram, a rootogram is a display of the distribution of some variable. In contrast to a histogram, the y-axis of a rootogram displays the square roots of the frequencies instead of the raw frequencies.

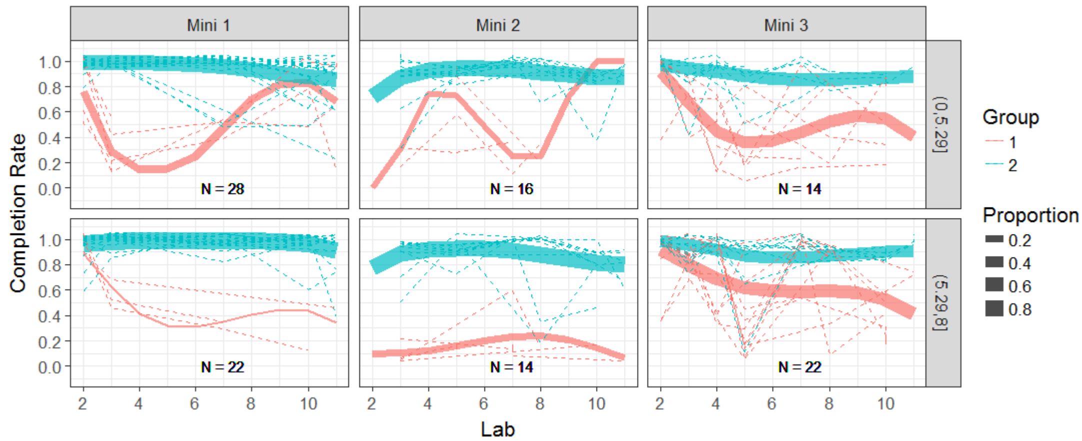

Figure 5: Completion rate trajectories, split by whether students answered more or fewer questions correctly on the quizzes than the average number (5.29) and in which mini they took the course. See the description of Figure 4 for more information. As we can see, the detected developmental patterns are not fundamentally different when we condition on the quiz scores instead of the final grades.

would follow the pre-determined path and approach the questions one after the other.

However, as Figure 6 shows, this does not seem to be the case: The figure depicts the time stamps (normalized to the range [0,1] for each user) for the first R programming exercise of three of the labs, where the user is prompted to type in and evaluate R code to answer the given question.[4]

As before, we have split the data displays such that students performing above the average on the final project are displayed in the lower frame, whereas the observations for the lower performing students are displayed in the upper frame.

Although the distribution of the relative frequency of the time stamps allows us to situate the exercises correctly at the beginning of the labs, it is still noteworthy to find that the students engaged with the task over the whole session. Some students revisited a problem after time has elapsed, while others started working on it at the end of the session. There is a great spread in the number of interactions by the individual students, both in the group of low and high performers. As previously discussed, this poses challenges concerning the validity of the number of clicks as a measure of learning engagement. Variation in click rates may largely reflect idiosyncrasies in how different students interact with their machines.

One can further observe that a complex question[5] (belonging to the plot on the right) results in a usage pattern characterized by students revisiting the problem over the entire session. In contrast, the two plots to the left display the user behavior with regards to a more basic question that may be solved with essentially a single interaction. Recalling that students from the first mini (whose clicks are colored in red in the plots) took the lab

[4]The labs in question are accessible under the following URLs:

- `http://isle.heinz.cmu.edu/94-842/lab04/`
- `http://isle.heinz.cmu.edu/94-842/lab08/`
- `http://isle.heinz.cmu.edu/94-842/lab11/`

[5]The question asked students to Write a function called calculateRowMeans that uses a for loop to calculate the row means of a matrix x., which involves programming concepts such as custom functions and control loops, unfamiliar for a multitude among the students.

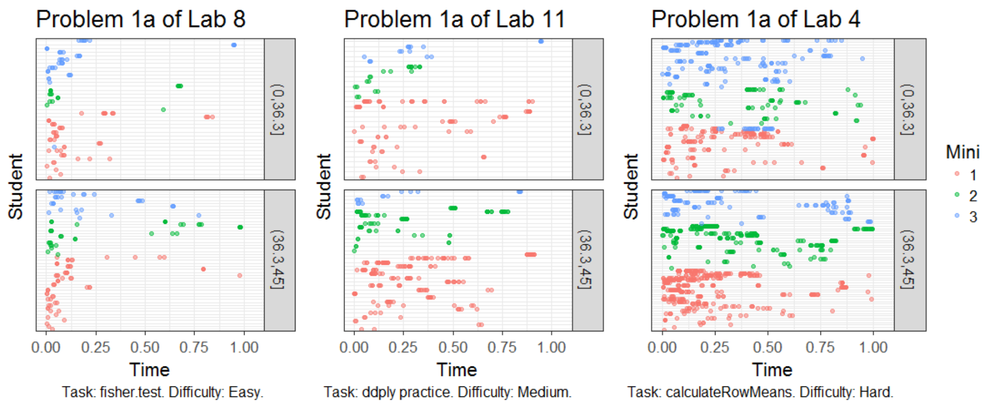

Figure 6: Student interactions with the R shells for the first problem from three of the labs (ordered by difficulty). Each dot represents a single action (Get Hint / Evaluate R Code / Show Solution), with the corresponding time in the session displayed on the x-axis, normalized to a number between zero and one (such that zero denotes the beginning of the session and one the end). Dots have been color-coded to signal which mini the respective student was enrolled in. In each plot, a line on the y-axis represents one students click stream, where for each mini students with higher number of clicks overall appear at the top and vice versa. As elaborated on in Section 5.2., it seems as if students do not proceed in a linear fashion through the labs, but revisit problems or even start them much later than anticipated. This is more pronounced for harder problems, which require more tries by the students.

sessions in class, the question arises whether this difference in setting has had any effect on the click rates. Judging from the two right-most plots belonging to Lab 4, it seems as if students who attempted the labs in class might have solved harder problems such as this one earlier than their fellow students, given that there are less clicks from them in the later parts of the session. While this has some intuitive appeal since students could ask teaching assistants for help, we hesitate to draw any firm conclusions at this time. This hypothesis merits further investigation, especially considering that the same pattern is not consistently observed over all questions.

8. Discussion

The analysis of student behavior confronts us with a plethora of unanticipated peculiarities that warrant further investigation. Analyses of user interactions with individual questions have demonstrated individualized engagement patterns that do not fit into the sequential order of the labs. A deeper understanding of this behavior will have implications for the structure of the lessons. Part of the ISLE framework is the ability to run experiments such as A/B tests and adapt material in a process of iterative design. In addition to this process of improving efficacy of our learning modules, we have demonstrated that modeling of student progress through trajectory analysis of one or more outcomes can be used to detect sub-groups of learners that follow a similar trajectory over time. Integrated into the ISLE system, this methodology can be used to create adaptive learning systems that are tailored to the peculiarities of a certain user group, for example by providing personalized emails to nudge students towards desired behavior. Preliminary experiments on nudging and personalized feedback conducted within ISLE suggest that this could be a successful approach to increasing engagement.

8.1 Limitations

Further research is needed to corroborate the findings of this study. Given that we had only access to a rather small student population in our pilot run, and that the ISLE framework has been under active development throughout, the experimental setting differed slightly over the various iterations of the course. For example, partway through we changed the design of some elements and started to collect data on certain actions that were previously unreported. Since completion of the ISLE labs was voluntary, we have missing data for quite a few students who did not spend much time with the labs. It is very likely that this missingness is not random. In this context, it would be of interest to better understand the differences among students who did or did not invest time in the labs. Finally, student populations varied over the different minis with respect to their majors and prior knowledge of statistics, so any cross-semester comparisons need to be taken with a grain of salt.

8.2 Future Work

In this paper, we have presented some preliminary findings of an analysis of data collected as part of the ISLE system. Since this project, which is currently in use in several courses at CMU, is still in active development, mining the gathered data is an ongoing challenge. We are currently working on facilities that make it easier to monitor and analyze the collected student data inside of the ISLE dashboard.

9. Conclusion

ISLE (*Interactive Statistics Learning Environment*) provides a framework for building e-learning lessons for statistics and enables instructors to track the learning trajectories of their students. Looking at student engagement in a course on R programming via group-based trajectory modeling, we found distinct student groups with characteristic temporal developments. An analysis of click rates showed that students have completed the created lessons in a non-linear manner, often delaying or revisiting problems during a session, an effect much more pronounced for harder questions. These observations demonstrate that students do not always act like one might expect. Further research is needed to validate and elucidate these findings, which could impact the construction and design of e-learning lessons for statistics instruction.

10. Acknowledgements

Seed funding for the project was provided through an internal ProSEED/Simon Initiative grant from Carnegie Mellon University. Our thanks to the support from Chad Hershock from the Eberly Center at Carnegie Mellon University, who assisted us in the design of the interactive labs.

References

1. Hofstein, A. & Lunetta, V. N. The laboratory in science education: Foundations for the twenty-first century. *Sci. Educ.* **88,** 28–54 (2004).
2. Tobin, K. Research on science laboratory activities: In pursuit of better questions and answers to improve learning. *Sch. Sci. Math.* **90,** 403–418 (1990).
3. Naps, T. L. *et al. Exploring the role of visualization and engagement in computer science education* in *ACM Sigcse Bull.* **35** (2002), 131–152.

4. Titterton, N. & Clancy, M. J. *Adding some lab time is good, adding more must be better: the benefits and barriers to lab-centric courses.* in *FECS* (2007), 363–367.
5. Pfaff, T. J. & Weinberg, A. Do hands-on activities increase student understanding?: A case study. *J. Stat. Educ.* **17,** 1–34 (2009).
6. Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A. & Bier, N. L. *Learning is Not a Spectator Sport* in *Proc. Second ACM Conf. Learn. @ Scale - L@S '15* (2015), 111–120. ISBN: 9781450334112. doi:10.1145/2724660.2724681. <http://dl.acm.org/citation.cfm?doid=2724660.2724681>.
7. Ma, W., Adesope, O., Nesbit, J. & Liu, Q. Intelligent Tutoring Systems and Learning Outcomes: A Meta-Analysis. *J. Educ. Psychol.* **106,** 901–918. ISSN: 1939-2176 (2014).
8. Means, B., Toyama, Y., Murphy, R. & Bakia, M. Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies. *US Dep. Educ.* <http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf> (2010).
9. Clark, R. C. & Mayer, R. E. *E-Learning and the Science of Instruction* ISBN: 9780787986834. doi:10.1002/9781118255971. arXiv: 1011.1669 (2008).
10. Fox, A. From MOOCs to SPOCs. *Commun. ACM* **56,** 38–40. ISSN: 00010782 (2013).
11. Ooms, J. The OpenCPU System: Towards a Universal Interface for Scientific Computing through Separation of Concerns. arXiv: 1406.4806. <http://arxiv.org/abs/1406.4806> (June 2014).
12. Nagin, D. S. Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychol. Methods* **4,** 139–157. ISSN: 1082-989X (1999).
13. Wedel, M. & DeSarbo, W. S. A mixture likelihood approach for generalized linear models. *J. Classif.* **12,** 21–55. ISSN: 01764268 (1995).