6.00x Introduction to Computer Science and Programming
MITx on edX – 2013 Spring

“6.00x is an introduction to computer science as a tool to solve real-world analytical problems.” – edx.org

Introduction. 6.00x is an introduction to using computation to solve real problems. The course is aimed at students with little or no prior programming experience who have a desire (or need) to understand computational approaches to problem solving. Some of the people taking the course will use it as a stepping-stone to more advanced computer science courses, but for many, it will be their first and last computer science course.

This was the second run of 6.00x, offered as an MITx on edX in the Spring of 2013, led by Professor Larry Rudolph (visiting MIT). The course had previously been run by Professors Eric Grimson and John Guttag, and Dr. Chris Terman.

Many participants in MITx on edX courses come with different intentions and levels of interest and many resemble shoppers and auditors in their course activity. This report identifies four subpopulations of participants (numbers below): those who registered, viewed some course content, explored the course, and were certified¹.

<table>
<thead>
<tr>
<th>Important Dates</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 20, 2012</td>
<td>Registered 72,920</td>
</tr>
<tr>
<td>Registration Open</td>
<td>Viewed 69,218</td>
</tr>
<tr>
<td>February 4, 2013</td>
<td>Course Launch</td>
</tr>
<tr>
<td>Course Close</td>
<td>Explored 5,744</td>
</tr>
<tr>
<td>June 4, 2013</td>
<td>Course Close</td>
</tr>
<tr>
<td>September 8, 2013</td>
<td>Date of Report</td>
</tr>
<tr>
<td>Date of Report Data</td>
<td>Certified 3,305</td>
</tr>
</tbody>
</table>

¹ See page 4 for definitions

This report was produced jointly by the MIT Office of Digital Learning, and the HarvardX Research Committee.

Course structure refers to the type, frequency, and order of resources in a given course. Within a MITx course on edX\(^2\), course structure is composed of a few base resource types (problems, videos, html pages), each categorized under a specific course component. Figure 1 visualizes the 575 resources of 6.00x in terms of their course component categories, where each line indicates a separate resource, and the length approximates the weight of the component toward final course grade (lecture sequences do not count). Text labels indicate selected chapters, which are the highest level on the “courseware” menu housing course content.

For 6.00x, Lecture Sequences formed the base learning material released in each chapter. Each sequence consisted of interview style and tablet annotation videos (orange) by Professors Grimson and Guttag, with finger exercises (black) interspersed between nearly all videos, some of which were code submissions. Graded assessment was administered through a combination of Homework (Silver) and Examinations (Red). Homework included numerical and formula response, and a large portion of problems were python-based code response questions. Supplemental learning components not pictured in Figure 1 included a threaded discussion forum for students and staff, and a staff and student editable wiki. A textbook was suggested, but not required.

The release of chapters in 6.00x occurred over 15 calendar weeks. Final course grades were determined by performance on lecture sequence exercises (10%), homework sets (40%), two midterms (25%), and a final exam (25%). Certification was granted to students whose final grades 55% or greater.

\(^2\) Often colloquially referred to as a Massive Open Online Course (MOOC)
To date, MITx courses on edX have had similar schedules of regularly released resources and periodic due dates. Like 6.00x, courses have used quantitative assessment supported by lecture sequences and textual material, as well as a few special learning components (see end of this report). In the figures that follow, we compare 6.00x data with those from other MITx courses on edX, from the Fall of 2012 to the Summer of 2013.

**Demographics.** For 6.00x, we have records of 72,290 students who registered for the course by the cutoff date of this report, September 8, 2013. According to self-reported demographics, registrants were 82.3% men, and as a group, highly educated: 38.3% have earned bachelor’s degrees, 19.9% have masters or professional degrees, and 3.2% of registrants had earned a doctorate (for certified students, percentages are 34.1%, 31.3%, and 5.7%, respectively). Distributions of level of education and gender are plotted in Figure 2 below, along with the average distributions for all MITx courses\(^3\). Note, roughly 7% of 6.00x registrants did not self-report one or more of the demographic categories.

---

**Figure 2:** Distributions of demographic variables collected at registration for all registrants reporting data, with MITx-wide comparisons for reference: (top left) education attainment, (top-right) gender, (bottom left) age, and (bottom right) geolocation via IP look-up.

---

As with other MITx courses on edX, 6.00x students came from all over the world. Based on geolocation of student IP addresses, 27.8% of registrants came from the United States, but many students also came from over 150 other countries around the world. Consistent with other MITx courses on edX, the median age of registrants and certified students were higher than those of typical MIT undergraduates; namely, 26 or 29 years, respectively. Distributions of geolocated country and age are plotted in Figure 2 for 6.00x, along side the average distributions for all MITx courses.

**Participants in 6.00x.** Who are the students taking this course, and how much did they participate? The degree and kind of participation in the course varied considerably among those who signed up for 6.00x. Furthermore, like many MOOCs, this course remained open to new registrants, so that enrollment continued to rise through the semester. This asynchronicity is a key feature of open online courses. One feature of low barriers to registration, for example, is a large number of students who registered but never actually viewed any course material.

To illustrate the considerable variability in participants by their actions, we identify four subpopulations of interest within this course: those who (1) only registered, and never accessed the courseware, (2) only viewed, i.e. non-certified registrants who accessed at least one chapter but less than half of chapters, (3) only explored, i.e. non-certified registrants who accessed half or more of the chapters, and those who were (4) certified, by earning a final grade at or higher than the cutoff.

Certificate earners are the most conventional subpopulation, defined by crossing a certain threshold of achievement on quantitative assessments (well defined for 6.00x). The “only explored” subpopulation seeks to represent students having interacted with a considerable amount of course material, while the “only viewed” subpopulation highlights participants having only sampled the course. Both the “only explored” and “only viewed” subpopulations contain participants with highly varying levels of activity (as highlighted in the following two sections on persistence and activity). The fourth sub-population of "only registered" contains users that registered, but never accessed the courseware. This is a substantial number in 6.00x and many of the MITx courses on edX.
Figure 4: Scatter plot of grade versus chapters viewed (left), highlighting student sub-populations; certified students are red points and all points are jittered. Histograms of grades and number of chapters viewed (right) distinguished by student certification status.

For 6.00x, the numbers of students who only registered, only viewed, only explored, or were certified, are shown in Fig. 3. The categories are defined as disjoint sets by using the “only” terminology, but the figure highlights the nested nature of registration, viewing, exploring, and certification. Students who were certified certainly viewed the course, but not all viewers were certified. Similarly, “explored and certified” are both subsets of viewed.

Activity in 6.00x. The activity of students in MITx courses on edX is logged; activity level can provide perspective into which courseware components are of interest to students, and which specific activities students are undertaking in a course, e.g. to earn certification.

How diverse is the activity of students in the course? Consider the grade earned by students, which is determined in 6.00x by quantitative assessment with regular due dates and periodic exams. A scatter plot (Figure 4) depicting earned grade versus the number of chapters visited, for each registered student,

Figure 5: Distribution of clicks (total course interactions) and active days, distinguished by certificate status.
illustrates the distinctions between the sub-populations of students who only viewed, only explored, or became certified in the course. This plot also illuminates some possible registrant types: “completionists” (perfect grade, viewed all chapters), “optimizers” (occasional registrants that view minimal chapters but earn a certificate), and “listeners” (viewed all chapters, zero grade – submitted no work for credit).

For 6.00x, edX recorded 55,047,961 events in its tracking logs, relating to students interacting with the course contents, through the cutoff date of this report. These events included interactions indicating when students played videos, attempted problems, browsed through text pages, read or posted entries in the forum, and other activities. Fig. 5 gives a glimpse into these data, presenting a histogram of the total number of events (“clicks”) recorded for all registrants with at least one event, and highlighting the population of students who earned a certificate in comparison with all other students. The figure also presents a histogram of the number of active days, where “active days” represents any day where a course interaction was logged. The distribution of active days for certificate earners is broad, further reflecting the variety of approaches by which a participant can earn a certificate.

![Unique Users](image)

**Figure 6:** Daily number of unique participants accessing courseware or course-site (pre-launch) via our described participant categories. Vertical dashed lines indicate the start and end of the course. Due to data loss, some information about user activity between January 21 and Feb 13 is missing and not included in the plot.

**Persistence in 6.00x.** How does the activity of students change as the course progresses, particularly with students who persist to earn certification? Figure 6 highlights courseware and course-site (pre-launch) interactions for three of the disjoint participant types: only viewed, only explored, and certified. In 6.00x, there is a noticeable periodicity in each curve that is highly correlated with content release and due dates. Certified students show a relatively stable weekly periodicity, while the periodicity of the viewed and explored curves disappears after about one-third and two-thirds of the course, respectively. It is important to

![Certification Rate](image)

**Figure 7:** Certification fraction versus registration week relative to course launch. Bubble size represents the number of unique users registering each week.
note that these curves do not account for registration date or for large times between user interactions. However, they do provide insight into the collective behavior of our defined populations. It is particularly interesting to formulate hypotheses about behavior early in the course, e.g., attempting to target struggling students that may benefit from intervention.

Do students who register early have a higher likelihood of earning a certificate than students who register late? Figure 7 depicts the certification fraction versus registration week relative to course launch (bubbles indicate total enrollment each week). Across MITx courses, certification rates do not depend greatly on registration dates prior to launch, although some courses have higher registration rates very early as well as just prior to launch. For 6.00x, between 6% and 12% of users registering prior to launch earn a certificate regardless of enrollment week. After launch, the rate drops dramatically. Because of the highly structured due dates in 6.00x, it is reasonable to attribute most of the post-launch drop in certification rates to registrants who have missed due dates and find it difficult or impossible to catch up.

Special Features
For special features found in the 2013 Spring 6.00x course, please reference the Fall 2012 6.00x MITx Course Report.

Comparison with Previous Course Runs
This was the second run of the 6.00x MITx course on edX. Between this Spring 2013 run, and the first, Fall 2012 run, there were similarities and differences:

- **Course content and structure:** In the second iteration of 6.00x, the order and frequency of course resources went largely unchanged. Exceptions included the addition of a number of finger exercises to weekly lecture sequences, as well as new midterm and final exams, which were broken into smaller, more modular, questions. Future work might address the question of how the density of finger exercises and exam questions impacts student behavior and performance.

- **Participant subpopulations:** The number of “only registered” students in the Fall 2012 and Spring 2013 courses shrank from 29,172 (34.5% of registrants) to 3,700 (5.1%). On the other hand, the number of “only explored” students grew from 46,365 (54.9%) to 63,475 (87%), indicating substantial changes in how students interacted with the courseware. One possible origin of the decrease in “only registered” student subpopulation is the shorter registration period in 6.00x Spring 2013 (46 days) compared to that for first run (64 days).

- **Demographics:** The populations of students participating in the second iteration of 6.00x largely resembled those of students in the Fall of 2012, in terms of sampled statistics for gender, age, and level of education, as well as geolocated country of origin.

- **Participants and activity:** The median number of clicks made by students in the “only explored” or “certified” subpopulations were similar across the two runs: 6896.5 for Fall 2012 and 6256 for Spring 2013. This may have been because the course structure and demographics were largely the same between iterations of 6.00x.

---