MITx 15.071x – The Analytics Edge

March 3, 2015

1 WELCOME!

In the last decade, the amount of data available to organizations has reached unprecedented levels. Data is transforming business, social interactions, and the future of our society. In this course, you will learn how to use data and analytics to give an edge to your career and your life. We will examine real world examples of how analytics have been used to significantly improve a business or industry. These examples include Moneyball, eHarmony, the Framingham Heart Study, Twitter, IBM Watson, and Netflix. Through these examples and many more, we will teach you the following analytics methods: linear regression, logistic regression, trees, text analytics, clustering, visualization, and optimization. We will be using the statistical software R and a spreadsheet software to build models and work with data. The contents of this course are essentially the same as those of the corresponding MIT class (The Analytics Edge). It is a challenging class, but it will enable you to apply analytics to real-world applications.

2 PREREQUISITES

Basic mathematical knowledge (at a high school level). You should be familiar with concepts like mean, standard deviation, and scatterplots. Mathematical maturity and prior experience with programming will decrease the estimated effort required for this class, but are not necessary to succeed.

3 COURSE OVERVIEW

The course is organized by units. To keep pace with the class, you are expected to complete all the work by the due dates indicated. For individual due dates, see the course calendar or the left-hand side navigation in the Courseware tab. Weekly coursework includes:

- Interactive lecture sequences;
- Recitation;
- Homework.

The course will also have a competition and a final exam.

4 INTERACTIVE LECTURE SEQUENCES

The lectures are presented in interactive sequences of videos and quick questions, and are posted in the *Courseware* section of the website. Each sequence includes a succession of short video clips and online questions, arranged in a logical progression. Please take the time to watch each video and complete each question in the sequence they are provided. Answer-check mechanisms are provided in these questions, and they will contribute towards your grade. They are designed to "quickly" test your understanding of the lecture material. When we work in R or LibreOffice during the lecture videos, we encourage you to follow along.

5 RECITATIONS

The recitations are a succession of short video clips, and are presented by one of the course teaching assistants. The purpose of these recitations is to complement the material taught during the lecture sequences, and to help you feel more comfortable with the method introduced during the lecture sequences, and we encourage you to follow along in R or LibreOffice.

6 HOMEWORK

A variety of problems will be assigned as homework. They will be issued at the start of each unit, in the *Courseware* section under the corresponding unit.

All homework assignments are due on Mondays at 23:59 UTC. **Late submissions in any format will not be accepted.**

While collaboration is welcomed and encouraged, **you are not allowed to post full solutions on the Discussion Forum for this class or anywhere else**. You are also not allowed to share complete solutions with each other.

7 COMPETITION AND FINAL EXAM

A significant portion of your final grade in 15.071x will be determined by the competition (after Unit 6) and the final exam (during the last week of the course). The course calendar lists the competition and final exam deadlines. The homework assignments are critical to learning the material and for doing well on the competition and exam.

The competition will be run on Kaggle, a platform for predictive modeling and analytics competitions. We will provide you with a dataset and problem, and ask you to use all of the methods you have learned to build the best possible predictive model. Discussion is welcomed and encouraged for the competition, but the submissions should be individual work.

The final exam will be during the last week of the course. Once you view the final exam problems, you must work on your own until you have submitted all your work, and should not discuss the exam until the deadline for exam submissions is past. The exam

will be open book, and will test your knowledge on all topics learned in the course.

You are not allowed to post answers to the final exam problems. Collaboration of any form is strictly forbidden for the final exam.

8 DISCUSSION FORUM

We will provide a discussion forum on our website for all students of 15.071x. You may use these forums to discuss course concepts, problem solving approaches, interesting references, or other topics related to the course. You may use it to ask questions, but no solutions can be posted on the discussion forum or shared with any students at any time. If solutions are posted on the discussion forum, they will be removed and students will be reminded of this rule. Students who repeatedly post solutions on the forum may face more serious consequences. Please observe the appropriate online etiquette as outlined in the Forum Guidelines posted in the Course Handouts section of the Course Info tab. The forum is moderated by course staff.

9 GRADING

Grades will be based on the following weighting: evaluations 2%, quick questions 8%, homework 50%, competition 15%, and final exam 25%. You will need to get a total mark of 55% or higher to earn a certificate.
Homework will be graded based on the best eight out of nine individual unit grades. Therefore, one homework assignment may be missed without a grade penalty.

10 CERTIFICATION

Online learners who achieve a passing grade in 15.071x earn a certificate of mastery. These certificates will indicate that you have successfully completed the course, but will not include a specific grade. Certificates will be issued by edX under the name of MITx. For this course in Spring 2015, there will be two certificate options: an ID verified certificate, and a free honor code certificate. For more information, please see the Certificates & Credits section of the Student FAQ (https://www.edx.org/student-faq).