15.S41: Software Tools for Business Analytics
January 13-17, 1-4 pm in E52-164
3 units, P/D/F grading

Course Administrators:
- Scott Alessandro, Director, Sloan Undergraduate Education. salessan@mit.edu, E52-150
- Austin Ashe, Associate Director, Sloan Undergraduate Education. aaashe@mit.edu, E52-154
- Natalie Petit, Assistant Director, Sloan Undergraduate Education. npetit@mit.edu, E52-154

Course Summary:
Because of the "big data revolution," there is an ever-increasing need for techniques for analyzing data, developing mathematical models, and using these models to make informed decisions. To get started in this process, one needs a working knowledge of business analytic software tools.

The goal of this course is to provide students with a baseline knowledge of business analytics software tools that they can use in MIT courses, UROPs involving data analysis, and summer internships or jobs after graduation.

Course Attendance Policy:
15.000 is graded Pass/Fail. Attendance at all sessions is required in order to receive a Pass. Listeners are allowed for students who have to miss a session(s).

Course Assignments:
Before each session, there will be pre-work, which generally involves installing software or data packages for use in that session. It is expected that students will complete this pre-work installation before each session.

There will be a small homework assignment after each session to reinforce what was covered in that session. The homework is evaluated on effort, not accuracy. So if you try and learn and submit the assignments, you will get credit.

Session Schedule:
Session 1 (Terminal and Github) – Monday, January 13, 1-4 pm, E52-164
Description: In this session we will give an overview to working with the terminal, Github, and a beginning introduction to the R programming language.

Session 2 (Data Wrangling and Visualization) – Tuesday, January 14, 1-4 pm, E52-164
Description: This session introduces basic techniques for data wrangling and visualization in R. Using contemporary best practices in statistical programming, we will explore a
powerful set of tools for efficiently preparing, analyzing, and visualizing complex data sets. The session does not require previous experience with R.

**Session 3 (Introduction to Machine Learning) – Wednesday, January 15, 1-4 pm, E52-164**
Description: This session introduces elementary methods for machine learning in R, focusing on the two classical supervised contexts of regression and classification. Building on our data preparation techniques from Session 2, we execute a complete pipeline from "raw data," to model training to model evaluation and communication. By the end of the session, students will creatively build and report on their own classification model.

**Session 4 (Case Study with R) – Thursday, January 16, 1-4 pm, E52-164**
Description: This session builds upon the experience in programming with R from previous sessions and introduces advanced concepts as a preview of the types of questions and analyses one can do with a full semester of R knowledge.

**Session 5 (Julia/JuMP) – Friday, January 17, 1-4 pm, E52-164**
Description: This session introduces the programming language "Julia" and the "JuMP" library. Julia is a high-level, high-performance dynamic programming language for technical computing, and JuMP is a library that allows us to easily formulate optimization problems and solve them using a variety of solvers. We will see how Julia and JuMP can be applied to solving real-world problems in operations and analytics.