15.S41: Software Tools for Business Analytics

January 14-18, 1-4 pm in E52-164
3 units, P/D/F grading
No Pre-requisites and No Prior Experience Needed

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.

Questions?: Scott Alessandro, salessan@mit.edu

Session 1 (Terminal and Github) – Monday, January 14, 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 15, 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. Our working example is a set of publicly available data from AirBnB. By the end of the session, students will construct a simple business intelligence dashboard for informing a realistic decision process. The session does not require previous experience with R.

Session 3 (Advanced R) – Wednesday, January 16, 1-4 pm, E52-164
Description: This session builds upon the experience in programming with R from Sessions 1 and 2 by introducing advanced concepts to create a bridge to Session 4.

Session 4 (Introduction to Machine Learning) – Thursday, January 17, 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. This session requires the material from Session 2 and 3.

Session 5 (JuMP/Julia) – Friday, January 18, 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.

Course 15 - Major, Minor, Take a Class