On Campus

Collaboratory Courses Demystify Big Data for Non-Specialists

Spring 2017



Matthew Jones (left) and Chris Wiggins co-teach a data-analytics course for humanities students. Photo by John Pinderhughes.

The availability of massive amounts of electronic data is transforming nearly every academic discipline and professional field today.

To prepare students to succeed in this data-rich world, Columbia has introduced a series of courses that combine instruction in disciplines such as business, journalism, public policy, history, and architecture with training in computer-based analytic techniques that are now commonly used in these fields. The eleven new courses, which launched this academic year, are co-taught by Columbia data scientists and faculty experts in each discipline.

“The goal is to give students the knowledge they need to be discriminating consumers and purveyors of statistical data and to possibly work together with data scientists,” says Patricia Culligan, a professor of civil engineering and associate director of the Columbia Data Science Institute who oversaw the formation of the new courses.

The courses were created with support from the Collaboratory@Columbia, a joint initiative of the Columbia Data Science Institute and Columbia Entrepreneurship that promotes digital literacy among students. Last year, the Collaboratory invited faculty from across the University to submit proposals for courses that would teach data-analysis techniques to a wide range of students; seven faculty teams were awarded fellowships to cover the costs of designing one or more new interdisciplinary courses. The initiative will support the development of at least twelve more courses over the next two years.

“The Collaboratory is operating in a grassroots manner by engaging faculty to incorporate data and computer-science skills into mainstream pedagogy,” says Richard E. Witten ’75CC, a University Trustee emeritus who directs Columbia Entrepreneurship and who led the creation of the Collaboratory. “This is generating a lot of radically fresh ideas.”

In a new course taught by public-policy expert Gregory Falco ’12SPS and computer scientist Adam Cannon, graduate students of public affairs and related subjects are learning to design more effective energy and environmental policies by using computer programs that can model their potential impact. And in a course taught by historian Matthew Jones and physicist Chris Wiggins ’93CC, undergraduate humanities and engineering majors are getting lessons in operating computational tools that scientists use to predict everything from election results to disease outbreaks.

“To be an informed citizen in the twenty-first century, you really need to be able to distinguish between crummy statistics and those that are truly explanatory or predictive,” says Jones. “By bringing students behind the scenes and showing them how different types of statistics get generated, we’re helping them make those determinations.”

Witten points out that the new courses supplement the University’s existing portfolio of interdisciplinary data-science offerings; more technically focused data-science

courses are available to Columbia students enrolled in degree and certificate programs that focus specifically on training journalists, economists, social scientists, and others in big-data techniques. But he suspects that more and more Columbia students will likely seek out at least some training in big-data analysis in the years ahead.

“It’s simply going to be seen as a part of a well-rounded education,” he says.



Guide to school abbreviations

All categories >