Honolulu Community College  
Topic 1: Predictive Analytics

The term "predictive analytics" basically refers to the use of quantitative methods to analyze past or real-time data in order to anticipate, change, or adapt to future situations and behavior. In higher education, and especially in the community colleges, the focus of analytics has generally been on retention and completion.

For the most part, the methods and focus of predictive analytics are not new to higher education.¹ What is new, and what has turned it into a must-have for colleges recently, is the dramatic expansion (and concurrent reduction in cost) of easily accessible data. With little more than SPSS or SAS—or even Excel—and several gigabytes of data, any school can easily generate interesting analyses that will likely yield insights, surprises and ideas for new tactics and strategies.

The more traditional forms of analytics are generally based on data generated by student application, registration, and course performance. The Predictive Analytics Reporting (PAR) Framework, a project of the Western Interstate Commission for Higher Education (WICHE), is an ambitious initiative designed to build and make available a structured, comprehensive set of data generated by institutions across the United States. Currently, this data is generated by 1.7 million student records and 8.1 million course records.

Institutions that are pursuing a more aggressive approach to analytics supplement traditional data with information generated by behavior that can now be tracked online (during registration and by learning management systems, for example) and on campus (purchasing, use of facilities, etc.)

It is useful (and perhaps sobering) to briefly consider how much data a single student might generate for such an institution in addition to a demographic profile and academic history (prior to and after admission): daily online actions (everything on the LMS for each course, every click on the campus web and registration sites), what and where any purchase is made with a point card, visits to access-controlled or monitored facilities (library, gym, computer lab, dormitories), and so forth. It is not surprising that data nerds see vast untapped potential in the streams of data that students (and everybody else on campus) generate for a properly fitted campus. It is also not surprising that for everybody else, that untapped potential raises a host of questions about the collection, storage, and use of that data. More on this in a second.

The University of Hawaii—a founding member of the PAR Framework mentioned above—has to date stayed mainly within the confines of more traditional, regression-based analyses of demographic and academic data. Examples of findings from UH-based institutional research:
• A study of 2009 and 2010 incoming freshmen at UHM found that the strongest predictors of their reenrollment for a second year were permanent address (mainland students were at the highest risk of not returning) and high school GPA (higher GPA led to better retention).\textsuperscript{2}

• A study of 2010 and 2011 first-time UHCC freshmen found the strongest effects on retention were the number of credits taken (15+ credits led to higher retention), high school GPA (higher GPA led to lower retention!), and ethnicity.\textsuperscript{3}

Brief studies of HonCC students with data provided by our PPIR office provides an interesting comparison to this larger system study:

Looking at 2012 and 2013 entering freshmen at HonCC, we looked at age, sex, distance from campus, CTE/LA, credits taken (cumulative, current, and earned), and enrollment in developmental English and math. Of these, credits attempted, credits earned, and enrollment in development math were significant predictors of a student’s reenrollment in the following Fall semester:

• The more credits a student earns during the year, the more likely he/she will reenroll, as we would expect.
• The more credits a student takes during the year, however, the less likely he/she will reenroll (in contrast to the larger study noted above).
• Students enrolled in a developmental math course are less likely to reenroll.
• Student permanent address appears to predict retention, although in varying ways that have not been consistent.
• As might be expected, when financial aid is considered, it becomes the most important predictor of retention.

While these findings are interesting in and of themselves, the key is how we can make use of them to improve retention. The 2012 Manoa IRO study demonstrated one way, by assigning specific numerical risk profiles (risk of not reenrolling) to categories of freshmen. So, for example, a male student from California not working on campus and receiving low financial aid is determined to be at high risk of not returning. Such profiles can then be used by faculty and staff to identify and actively reach out to such students, in an effort to counteract those risk factors. This kind of analysis can be extended to all kinds of variables, provided that the data is available.

The development and use of risk profiles, of course, raises interesting ethical issues in how and when predictive analytics can be used appropriately.\textsuperscript{4} Such issues are even more pressing as colleges move towards more aggressive uses of analytics by using student behavior (as measured online and with ID/point card systems) to classify them, shape their available options, and trigger institutional intervention. It is important that a campus community understand which data is being collected and how its analysis is informing and shaping retention policy and practice.
As we look at analytics as a tool to improve retention and completion, we welcome input from the campus community on these and other questions:

- Should HonCC be using predictive analytics to “create” a cohort of students to reach out to?
- How can this information support our instructors in the classroom or change instructional/classroom management practices?
- How can this information change the way we provide academic and support services to our students?


3 Manoa Institutional Research Office, “Predictive Analysis on Freshmen Retention at University of Hawaii Community Colleges (Draft),” August 2014.