Data Science Courses (Statistics and Actuarial Science) (DATA)

DATA Courses
This is a list of courses with the subject code DATA. For more information, see Data Science and Statistics and Actuarial Science (College of Liberal Arts and Sciences) in the Catalog.

DATA:3120 Probability and Statistics 4 s.h.
Models, discrete and continuous random variables and their distributions, estimation of parameters, testing statistical hypotheses. Prerequisites: MATH:1560 or MATH:1860. Same as IGPI:3120, STAT:3120.

DATA:4750 Probabilistic Statistical Learning 3 s.h.
Essential machine learning and statistics ideas that are critical in analyzing modern complex and large data; supervised learning topics include linear models, deep neural networks, and nonparametric models; essential topics include nonlinear dimension reduction, clustering, and recommender systems. Prerequisites: (CS:1210 with a minimum grade of C- or ENGR:2730 with a minimum grade of C-) and (MATH:2700 or MATH:2550) and (STAT:2010 or STAT:2020 or STAT:4200) and STAT:4540.

DATA:4880 Data Science Creative Component 1 s.h.
Readings, group discussions, and short-term projects in area of data science; emphasis on communication of ideas learned in student's data science coursework, data ethics, and potential bias in algorithms.

DATA:4890 Data Science Practicum 2 s.h.
On- or off-campus internship or group-based consulting project that provides experience in a real-world setting; application of knowledge and techniques learned in coursework; practice in communicating results to others.

DATA:5890 M.S. Data Science Practicum 2 s.h.
On- or off-campus internship or group-based consulting project that provides experience in a real-world setting; application of knowledge and techniques learned in coursework and practice communicating results to others.

DATA:6200 Predictive Analytics 3 s.h.
Linear mixed models; generalized linear mixed models; generalized additive models; applications of these models using associated R packages. Prerequisites: STAT:4560. Corequisites: STAT:4561. Requirements: comfort working with R software environment. Same as ACTS:6200.

DATA:7350 High-Dimensional Probability for Data Science 3 s.h.
Nonasymptotic probability with a view towards applications in data science; concentration inequalities for functions of independent variables, martingale inequalities, entropy method, random matrices, matrix inequalities, suprema of random processes, and sparse recovery. Prerequisites: STAT:5101. Requirements: linear algebra course and familiarity with R or Python.