Data Science (DSCI)

Courses

DSCI 101. Foundations of Data Science I. 4 Credits.
This course utilizes a quantitative approach to explore fundamental concepts in data science. Students will develop key skills in programming and statistical inference as they interact with real-world data sets across a variety of domains. Ethical and privacy concerns are explored. Sequence with DSCI 102.
Additional Information:
Science Area

DSCI 102. Foundations of Data Science II. 4 Credits.
This course expands upon critical concepts and skills introduced in DSCI 101. Topics include the normal distribution, confidence intervals, regression, and classifiers. Sequence with DSCI 101.
Requisites: Prereq: DSCI 101, MATH 101 (or equivalent placement score) or any other college-level math course.

DSCI 196. Field Studies: [Topic]. 1-12 Credits.
Repeatable.
Repeatable 99 times

DSCI 198. Workshop: [Topic]. 1-12 Credits.
Repeatable.
Repeatable 99 times

DSCI 199. Special Studies: [Topic]. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 299. Special Studies: [Topic]. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 311. Principles and Techniques of Data Science. 4 Credits.
Intermediate and advanced techniques in data science. Topics include managing data using software programs, data cleaning, handling text, dimensionality, principle component analysis, regression, classification and inference.
Requisites: Prereq: DSCI 102, CS 211, MATH 342.

DSCI 345M. Probability and Statistics for Data Science. 4 Credits.
Introduction to probability and statistics, with an emphasis upon topics relevant for data science. Students cannot get credit for both MATH 343 and DSCI 345M/MATH 345M.
Requisites: Prereq: MATH 342, CS 211.
Equivalent to: MATH 343, MATH 345M

DSCI 350M. Humanities Research Data Management. 4 Credits.
This course provides students with theoretical and practical experience in collecting, processing, archiving, and publishing humanities data (images, video, sound, text, maps, etc.) gathered from galleries, libraries, archives, and museums (GLAMs). Multilisted with LIB 350M.
Equivalent to: LIB 350M

DSCI 372M. Machine Learning for Data Science. 4 Credits.
Introduction to Machine Learning, with an emphasis on topics relevant for data science. Multilisted with CS 372M.
Requisites: Prereq: CS 212, DSCI 345M, MATH 342.
Equivalent to: CS 372M

DSCI 399. Special Studies: [Topic]. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 400M. Temporary Multilisted Course. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 401. Research: [Topic]. 1-12 Credits.
Repeatable.
Repeatable 99 times

DSCI 402. Supervised College Teaching. 1-6 Credits.
Repeatable for a max of 6 credits.
Repeatable 99 times

DSCI 403. Thesis. 1-12 Credits.
Repeatable.
Repeatable 99 times

DSCI 404. Internship: [Topic]. 1-12 Credits.
Repeatable.
Repeatable 99 times

DSCI 405. Reading and Conference: [Topic]. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 406. Field Studies: [Topic]. 1-12 Credits.
Repeatable.
Repeatable 99 times

DSCI 407. Seminar: [Topic]. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 409. Terminal Project. 1-12 Credits.
Repeatable.
Repeatable 99 times

DSCI 410. Experimental Course: [Topic]. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 410L. Experimental Course: [Topic]. 4 Credits.
Repeatable.
Repeatable 99 times

DSCI 411. Capstone Project. 4 Credits.
This course for Data Science majors provides a student the opportunity to apply the theoretical knowledge and techniques acquired during the Data Science degree curriculum to a project involving real data from the student's domain of specialization. Requires an average 3.75 GPA in courses required.
Requisites: Prereq: DSCI 311, DSCI 372M, PHIL 223.

DSCI 601. Research: [Topic]. 1-16 Credits.
Repeatable.
Repeatable 99 times

DSCI 604. Internship: [Topic]. 1-9 Credits.
Repeatable.
Repeatable 99 times

DSCI 610. Experimental Course: [Topic]. 1-5 Credits.
Repeatable.
Repeatable 99 times

DSCI 625. Foundational Mathematics for Data Science. 4 Credits.
Key math for data science models, algorithms, and processes. Taught at a level accessible to those with minimal college-level math. Topics include differential calculus including higher-order and partial derivatives, linear algebra including linear transformations and eigenanalysis, and select topics in discrete mathematics.

DSCI 626. Foundational Statistics for Data Science. 4 Credits.
An introduction to data visualization and statistical inference. Taught at a level accessible to those with no background in statistics. Topics include basic simulation, publication-quality plot production, parameter estimation, hypothesis testing and p-values, linear models, and non-parametric statistical analysis.

DSCI 627. Foundational Programming for Data Science. 4 Credits.
An introduction to computer programming for data science. Taught at a level accessible for those with no prior programming experience. Topics include an introduction to programming, debugging, reproducible code and documentation, data cleaning and visualization, libraries, and a basic introduction to simple machine-learning approaches.

DSCI 631. Data Access and Management. 4 Credits.
Data management using structured and unstructured databases. Students learn to interface with common database systems. Working with structured databases, students learn the fundamentals of relational data and query languages. Students also learn to work with unstructured databases.

DSCI 632. Statistics for Data Science. 4 Credits.
Probability for classical and Bayesian inference. This course will cover statistical uncertainty as applied to hypothesis testing, experimental design, and parameter estimation. Students will gain exposure to common probability distributions in the Sciences, and their application in statistical models, validation of model assumptions, and hypothesis formulation and testing. Basic maximum likelihood methods using linear models (regression, ANOVA) will be covered. Will include an introduction to Bayesian statistics.

DSCI 633. Machine Learning I. 4 Credits.
Applied concepts and methods in machine learning. This course will cover fundamental concepts, algorithms, and techniques. High-level conceptual and statistical tradeoffs in machine learning will feature prominently, as will key processes in model optimization, including resampling methods and an introduction to gradient descent. Both supervised and unsupervised learning techniques covered.

DSCI 634. Machine Learning II. 4 Credits.
Advanced methods in machine learning. This course will feature a diverse range of techniques to round out student mastery of cutting-edge predictive modeling. Includes clustering and multivariate ordination redux, Bayesian model fitting procedures with Hamiltonian Monte Carlo, and deep learning with various classes of artificial neural network. Students will also gain exposure to the art of feature engineering and learn how to define effective, tailored loss functions. Sequence with DSCI 633.

DSCI 635. Data Mining, Exploration, and Visualization. 4 Credits.
Effective data summarization and description. This course will teach students how to efficiently develop intuition for a dataset. Best practices in data manipulation, including key concepts in data cleaning and tidying will be covered. Students will learn how to summarize data with summary statistics and ordination techniques, as well as visualize data using a range of transformations and plotting approaches.
Requisites: Prereq: DSCI 631.

DSCI 636. Synthetic Project Capstone. 4 Credits.
Project-driven application of data science knowledge and skills. Students will gather, tidy, visualize, and analyze a large data set conducive to a predictive modeling problem. Curated data sets from biological and materials sciences will be provided by program faculty and colleagues. Key learning outcomes include a reproducible workflow published in a code repository, original design and implementation of a machine learning model, and communications of finding in a written report.
Requisites: Prereq: DSCI 631, DSCI 632, DSCI 633, DSCI 634, DSCI 635, PHIL 623.