DATA SCIENCE (DS)

Courses primarily for undergraduates:

DS 1100: Orientation to Data Science
Credits: Required. Contact Hours: Lecture 1.
Introduction to the procedures and policies of Iowa State University and the Data Science program, test-outs, honorary societies, etc. Issues relevant to student adjustment to college life will also be discussed. Offered on a satisfactory-fail basis only. (Typically Offered: Fall)

DS 2010: Introduction to Data Science
Credits: 3. Contact Hours: Lecture 2, Laboratory 2.
Prereq: Satisfactory math placement test score (ALEKS, 51+)
Data Science concepts and their applications; domain case studies with applications in various fields; overview of data analysis; major components of data analysis pipelines; computing concepts for data science; descriptive data analysis; hands-on data analysis experience; communicating findings to stakeholders, and ethical issues in data science. Placement scores can be found at: https://math.iastate.edu/academics/undergraduate/aleks/placement/. Typically Offered: Fall (annually), Spring (annually), Summer (irregularly).

DS 2020: Data Acquisition and Exploratory Data Analysis
Credits: 3. Contact Hours: Lecture 3.
Prereq: DS 2010
Data acquisition: file structures, web-scraping, database access; ethical aspects of data acquisition; types of data displays; numerical and visual summaries of data; pipelines for data analysis: filtering, transformation, aggregation, visualization and (simple) modeling; good practices of displaying data; data exploration cycle; graphics as tools of data exploration; strategies and techniques for data visualizations; basics of reproducibility and repeatability; web-based interactive applets for visual presentation of data and results. Programming exercises. (Typically Offered: Fall, Spring)

DS 3010: Applied Data Modeling and Predictive Analysis
Credits: 3. Contact Hours: Lecture 3.
Prereq: DS 2020 and STAT 1010, STAT 1040*, STAT 1050,* STAT 2010*,* STAT 2260*,* STAT 2310*,* STAT 3050*,* STAT 3220 *or* STAT 3300
Elements of predictive analysis such as training and test sets; feature extraction; survey of algorithmic machine learning techniques, e.g. decision trees, Naive Bayes, and random forests; survey of data modeling techniques, e.g. linear model and regression analysis; assessment and diagnostics: overfitting, error rates, residual analysis, model assumptions checking; communicating findings to stakeholders in written, oral, verbal and electronic form, and ethical issues in data science. Participation in a multi-disciplinary team project. (Typically Offered: Fall, Spring)

DS 3030: Concepts and Applications of Machine Learning
Credits: 3. Contact Hours: Lecture 3.
Prereq: DS 2020*,* MATH 2070*,* MATH 2650*, and* STAT 3010
Machine learning concepts such as training and test sets; feature extraction; principles of machine learning techniques; regression; pattern recognition methods; unsupervised learning techniques; assessment and diagnostics: overfitting, error rates, residual analysis, model assumptions checking, feature selection; ethical issues in data science; communicating findings to stakeholders in written, oral, visual and electronic form. (Typically Offered: Fall)

DS 3900: Special Topics in Data Science
Credits: 1-3. Contact Hours: Lecture 3.
Repeatable, maximum of 6 credits.
Prereq: Instructor Permission for Course
Lecture/seminar on special topics in Data Science. (Typically Offered: Fall, Spring, Summer)

DS 4010: Data Science Capstone
Credits: 3. Contact Hours: Lecture 3.
Prereq: DS 3010 *or DS 3030*
Students work as individuals and teams to complete the planning, design, and implementation of a significant multi-disciplinary project in data science. Oral and written reports. Offered irregularly. (Typically Offered: Fall, Spring)

DS 4220X: Mathematical Principles of Data Science
(Cross-listed with MATH 4220X).
Credits: 3. Contact Hours: Lecture 3.
Mathematical foundations of algorithms in data science. Topics include Riemann-Stieltjes integration, Riesz-Markov theorem, Stone-Weierstrass theorem, Universal Approximation theorem, reproducing kernel Hilbert spaces, Cauchy and Fourier kernels, convergence of clustering algorithms, and topological persistence. (Typically Offered: Spring)