

# DATA SCIENCE-DS (DSCI)

## Courses

### DSCI 100 First Year Seminar in Data Science Credit: 1 (0-0-1)

**Course Description:** Introduction to problems and techniques in data science.

**Prerequisite:** None.

**Registration Information:** Freshman or sophomore Data Science majors only.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 235 Data Wrangling Credits: 2 (1-0-1)

**Course Description:** Introduce tools and techniques for handling, cleaning, extracting, and organizing data.

**Prerequisite:** (CS 150B with a minimum grade of C or CS 152 with a minimum grade of C) and (CS 165 with a minimum grade of C or STAT 341 with a minimum grade of C).

**Restriction:** Must be a: Undergraduate.

**Registration Information:** Must register for lecture and recitation. Sections may be offered: Online.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 320 Optimization Methods in Data Science Credits: 3 (3-0-0)

**Also Offered As:** MATH 320.

**Course Description:** Linear and non-linear programming, convex sets and functions, convex and non-convex optimization problems, duality, Newton's methods, barrier methods, linear equality and inequality constraints. Emphasis on computation methods and programming.

**Prerequisite:** (CS 163 or CS 164 or CS 165 or CS 220 or DSCI 235) and (DSCI 369 or MATH 369) and (MATH 255 or MATH 256 or MATH 261).

**Registration Information:** Credit not allowed for both DSCI 320 and MATH 320.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 335 Inferential Reasoning in Data Analysis Credits: 3 (3-0-0)

**Course Description:** The role of models in making inferences from data; estimation vs testing; frequentist vs. Bayesian interpretations of probability; interpretability of statistical summaries of data; misinterpretations and misuses of popular inferential methods; confounder, collider, mediator, and post-treatment variables in causal diagrams; use of simulation to investigate properties of model-based inference and consequences of assumption violations; recent controversies surrounding statistical inference.

**Prerequisite:** None.

**Registration Information:** (CO 300 or CO 301B or CO 301C or CO 302 or JTC 300) and [(STAT 201 or STAT 204) and STAT 302A] or STAT 301 or STAT 307 or STAT 315.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 336 Data Graphics and Visualization Credit: 1 (1-0-0)

**Course Description:** Data graphics and visualization techniques for data science.

**Prerequisite:** STAT 158.

**Registration Information:** This is a partial semester course.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 369 Linear Algebra for Data Science Credits: 4 (3-0-1)

**Course Description:** Techniques in linear algebra related to data science. Matrices, bases, subspaces, linear independence, dimension, change of basis, projections, linear systems of equations, least squares, matrix factorizations. Singular value decomposition, angles between subspaces.

**Prerequisite:** MATH 157 and MATH 159 or MATH 155 or MATH 156 or MATH 160 or MATH 161.

**Registration Information:** Must register for laboratory and recitation.

Sections may be offered: Online.

**Terms Offered:** Fall, Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 445 Statistical Machine Learning Credits: 3 (3-0-0)

**Course Description:** Algorithms and statistical methods for regression, classification, and clustering; hands-on experience in analyzing data and running machine learning experiments.

**Prerequisite:** (DSCI 369 or MATH 369) and (STAT 341, may be taken concurrently).

**Registration Information:** Sections may be offered: Online. Credit not allowed for both CS 345 and DSCI 445.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 473 Introduction to Geometric Data Analysis Credits: 2 (2-0-0)

**Course Description:** Geometric techniques for analyzing high-dimensional and complex data. Techniques for data reduction and analysis.

**Prerequisite:** DSCI 369.

**Registration Information:** This is a partial semester course.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

### DSCI 475 Topological Data Analysis Credits: 2 (2-0-0)

**Course Description:** Topological techniques for analyzing high-dimensional or complex data. Topics include clustering, dendrograms, a visual introduction to topology, data modeling and visualization, and selected topics from nonlinear dimensionality reduction, graph-based models of data, Reeb graphs, multi-scale approaches to data, and persistent homology.

**Prerequisite:** DSCI 369 or MATH 369.

**Registration Information:** This is a partial semester course.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 478 Capstone Group Project in Data Science Credits: 4 (0-0-8)**

**Course Description:** Group-project-based capstone, in which small groups of students from each Data Science degree concentration work collectively on a problem in data science.

**Prerequisite:** DSCI 445.

**Restriction:** Must be a: Undergraduate.

**Registration Information:** Senior standing only.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 498 Research Credits: Var[1-3] (0-0-0)**

**Course Description:**

**Prerequisite:** CS 150B.

**Restrictions:** Must not be a: Freshman. Must be a: Undergraduate.

**Registration Information:** Sophomore standing. Written consent of instructor. Enrollment in Data Science major.

**Terms Offered:** Fall, Spring, Summer.

**Grade Mode:** Instructor Option.

**Special Course Fee:** No.

**DSCI 510 Linux as a Computational Platform Credit: 1 (1-0-0)**

**Course Description:** Use of the Linux operating system for computational work using command-line tools; basic Linux commands, running and managing jobs, installing software.

**Prerequisite:** None.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. This is a partial semester course. Credit allowed for only one of the following: CS 580A4, DSCI 510, or NSCI 580A4.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 511 Genomics Data Analysis in Python Credits: 2 (1-0-1)**

**Course Description:** Analyzing complex data sets using Python.

**Prerequisite:** DSCI 510, may be taken concurrently.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Must register for lecture and recitation. This is a partial semester course. Credit not allowed for both DSCI 511 and NSCI 580A5.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 512 RNA-Sequencing Data Analysis Credit: 1 (0-2-0)**

**Course Description:** Hands-on experience with tools for analysis of next generation sequencing data.

**Prerequisite:** DSCI 510, may be taken concurrently.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. This is a partial semester course. Credit not allowed for both DSCI 512 and NSCI 580A3.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 520 Optimization for Data Science and AI Credits: 3 (3-0-0)**

**Course Description:** Mathematical foundations of optimization algorithms for Data Science and AI. Local, global minima, convex optimization, gradient and sub gradient methods, neural networks, backpropagation, trace optimization, Laplacian eigenmaps, duality theory, support vector machines, elastic net, semidefinite programming and applications.

**Prerequisite:** (CS 152) and (DSCI 369 or DSCI 569A or MATH 369 or MATH 569A) and (MATH 261).

**Restriction:** Must not be a: Freshman, Sophomore, Junior.

**Registration Information:** Senior standing. Sections may be offered: Online.

**Term Offered:** Fall.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 521 Introduction to Optimization for Data Science Credit: 1 (1-0-0)**

**Course Description:** The basic mathematical tools used in optimization algorithms. Decision variables, local, global minima, Taylor's theorem, Newton's method, steepest descent, neural networks, backpropagation, stochastic gradient descent, and introduction to PyTorch.

**Prerequisite:** (CS 152) and (DSCI 569A or MATH 569A) and (MATH 160).

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Sections may be offered: Online. Credit allowed for only one of the following: DSCI 521, DSCI 580A6, or MATH 580A6.

**Terms Offered:** Fall, Spring, Summer.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 525 Data Wrangling with Python Credits: 2 (1-0-1)**

**Course Description:** Python libraries and techniques for extracting, cleaning, visualizing, and analyzing data.

**Prerequisite:** CS 152.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Must register for lecture and recitation.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 535 Inferential Reasoning in Data Science Credits: 2 (2-0-0)**

**Course Description:** The role of models in making inferences from data; estimation vs testing; frequentist vs. Bayesian interpretations of probability; interpretability of statistical summaries of data; misinterpretations and misuses of popular inferential methods; confounder, collider, mediator, and post-treatment variables in causal diagrams; use of simulation to investigate properties of model-based inference and consequences of assumption violations.

**Prerequisite:** None.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Sections may be offered: Online.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 546 Machine Learning for Data Science Credits: 3 (3-0-0)**

**Course Description:** Concepts, methods, and tools for machine learning and data analysis using Python. Design and implementation of rigorous machine learning experiments; statistical analysis of multivariate data; regression and classification algorithms. Effective written communication of machine learning tasks, methods, and results.

**Prerequisite:** CS 152.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 550 Computational Physical Statistics Credits: 3 (3-0-0)**

**Also Offered As:** PH 550.

**Course Description:** Computational physical statistics, including analysis of large datasets, Bayesian and frequentist approaches to theoretical and experimental error analysis, sampling techniques in statistics and statistical/quantum physics, and Markov chain Monte Carlo. Example systems include classical spin systems and Euclidean-time formulation of quantum mechanics. Emphasis is placed on application through numerical methods.

**Prerequisite:** (MATH 256 or MATH 261) and (PH 142).

**Registration Information:** Credit not allowed for both DSCI 550 and PH 550.

**Term Offered:** Spring.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 569A Linear Algebra for Data Science: Matrices and Vectors Spaces Credit: 1 (1-0-0)**

**Also Offered As:** MATH 569A.

**Course Description:** A basic introduction to matrices and linear algebra with preparation to pursue further studies in the applications of matrices with an emphasis on the foundations of data science.

**Prerequisite:** MATH 124 and MATH 126 or MATH 127.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Offered as an online course only. Credit not allowed for both DSCI 569A and MATH 569A.

**Terms Offered:** Fall, Spring, Summer.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 569B Linear Algebra for Data Science: Geometric Techniques for Data Reduction Credit: 1 (1-0-0)**

**Also Offered As:** MATH 569B.

**Course Description:** Projections, data fitting and over-determined linear systems, eigenvectors and eigenvalues, the spectral theorem for symmetric matrices, data driven bases, principal component analysis, the singular value decomposition.

**Prerequisite:** DSCI 569A or MATH 569A.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Sections may be offered: Online. Credit allowed for only one of the following: DSCI 569B, MATH 569B, or MATH 580A3.

**Terms Offered:** Fall, Spring, Summer.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 569C Linear Algebra for Data Science: Matrix Factorizations and Transformations Credit: 1 (1-0-0)**

**Also Offered As:** MATH 569C.

**Course Description:** Advanced algorithms for the characterization of data using matrix factorizations and transformations.

**Prerequisite:** DSCI 569B or MATH 569B.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Sections may be offered: Online. Credit allowed for only one of the following: DSCI 569C, MATH 569C, or MATH 580A4.

**Terms Offered:** Fall, Spring, Summer.

**Grade Mode:** Traditional.

**Special Course Fee:** No.

**DSCI 569D Linear Algebra for Data Science: Theoretical Foundations Credit: 1 (1-0-0)**

**Also Offered As:** MATH 569D.

**Course Description:** Theoretical development of linear algebraic tools for data science; theorem and proof driven.

**Prerequisite:** DSCI 569C or MATH 569C.

**Restriction:** Must be a: Graduate.

**Registration Information:** Graduate standing. Sections may be offered: Online. Credit allowed for only one of the following: DSCI 569D, MATH 569D, or MATH 580A5.

**Terms Offered:** Fall, Spring, Summer.

**Grade Mode:** Traditional.

**Special Course Fee:** No.