SDS - Statistics and Data Sciences

Statistics and Data Sciences: SDS

Lower-Division Courses

SDS 301 (TCCN: MATH 1342). Elementary Statistical Methods.

Covers the fundamental procedures for data organization and analysis. Subjects include frequency distributions, graphical presentation, sampling, experimental design, inference, and regression. Three lecture hours a week for one semester. Only one of the following may be counted: Educational Psychology 308, Statistics 309, 309H or Statistics and Data Sciences 301.

SDS 302F. Foundations of Data Analysis.

Introduction to data analysis and statistical methods. Subjects include random sampling; principles of observational study and experimental design; data summaries and graphics; and statistical models and inference, including the simple linear regression model and one-way analysis of variance. Three lecture hours and one laboratory hour a week for one semester. Only one of the following may be counted: Statistics and Data Sciences 302, 302F, 306.

SDS 110T, 210T, 310T, 410T. Topics in Statistics and Computation.

For each semester hour of credit earned, one lecture hour a week for one semester. Statistics and Data Sciences 110T and Statistics and Scientific Computation 110T may not both be counted. May be repeated for credit when the topics vary.

SDS 311C. Numbering Race.

Same as African and African Diaspora Studies 302M. Subjects include conceptualization and operationalization in quantitative measurement, the calculation and interpretation of descriptive statistics and statistical relationships, the application of statistical techniques to understand social phenomenon, and techniques for presenting results from quantitative analysis. Three lecture hours a week for one semester. Only one of the following may be counted: African and African Diaspora Studies 302M, 317D (Topic: Numbering Race), Statistics and Data Sciences 310T (Topic: Numbering Race), 311C.

SDS 313. Introduction to Data Science.

Introduction to the principles and practice of data science. Explore R and reproducible data analysis; summarizing data using descriptive statistics; data visualization and storytelling; data wrangling and relational data; basic prediction and classification using regression models; and programming in R. The equivalent of three lecture hours a week for one semester. Only one of the following may be counted: Statistics and Data Sciences 313, 322E, 348.

SDS 315. Statistical Thinking.

Introduction to the fundamental ideas of statistical thinking with R programming. Explore survey, experimental, and observational study design; common sources of random and systematic error in data; the bootstrap as a tool for quantifying uncertainty; hypothesis testing; regression; and the role of statistics in an ethical and just society. The equivalent of three lecture hours a week for one semester. Prerequisite: Statistics and Data Sciences 313 with a grade of at least C-.

SDS 318. Introduction to Statistical and Scientific Computation.

An introduction to quantitative analysis using fundamental concepts in statistics and scientific computation. Includes probability, distributions, sampling, interpolation, iteration, recursion, and visualization. Three lecture hours and one laboratory hour a week for one semester. Statistics and Data Sciences 318 and Statistics and Scientific Computation 318 may not both be counted.

Upper-Division Courses

SDS 320E. Elements of Statistics.

Introduction to statistics. Subjects include probability; principles of observational study and experimental design; statistical models and inference, including the multiple linear regression model and one-way analysis of variance. R programming is introduced. Three lecture hours and one laboratory hour a week for one semester. Only one of the following may be counted: Statistics and Data Sciences 320E, 320H, and 328M.

SDS 320H. Elements of Statistics: Honors.

Introduction to statistics. Subjects include probability; principles of observational study and experimental design; statistical models and inference, including the multiple linear regression model and one-way analysis of variance. R programming is introduced. Three lecture hours and one laboratory hour a week for one semester. Only one of the following may be counted: Statistics and Data Sciences 320E, 320H, and 328M.

SDS 321. Introduction to Probability and Statistics.

Covers fundamentals of probability, combinatorics, discrete and continuous random variables, jointly distributed random variables, and limit theorems. Using probability to introduce fundamentals of statistics, including Bayesian and classical inference. The equivalent of four lectures hours a week. Statistics and Data Sciences 321 and 431 may not both be counted. Prerequisite: Mathematics 408C, 408L, 408R, 408S, or 408Q with a grade of at least C-.

SDS 222, 322. Introduction to Scientific Programming.

Introduction to programming using both the C and Fortran (95/2003) languages, with applications to basic scientific problems. Covers common data types and structures, control structures, algorithms, performance measurement, and interoperability. For each semester hour of credit earned, one lecture hour a week for one semester. Only one of the following may be counted: Statistics and Data Sciences 222, 322, 292, 392, Statistics and Scientific Computation 222, 322, 292, 392. Prerequisite: Credit or registration for Mathematics 408C, 408K, or 408N.

SDS 322E. Elements of Data Science.

Explore data science tools and examine data wrangling; exploratory data analysis and data visualization; markdown and data workflow; simulation-based inference; and classification methods. R programming is emphasized and Python programming is introduced. Three lecture hours and one laboratory hour a week for one semester. Only one of the following may be counted: Statistics and Data Sciences 313, 322E, 348. Prerequisite: An introductory statistics course.

SDS 323. Statistical Learning and Inference.

An introduction to statistical influence, broadly construed as the process of drawing conclusions from data, and to quantifying uncertainty about said conclusions. Covers the major schools of thought that influence modern scientific practice, including classical frequentist methods, machine learning, and Bayesian inference. Three lecture hours a week for one semester. Statistics and Data Sciences 323 and Statistics and Scientific Computation 323 may not both be counted.

Prerequisite: Statistics and Data Sciences 321 (or Statistical and Scientific Computation 321) or the equivalent.

SDS 324E. Elements of Regression Analysis.

Explore the use of regression analysis in applied research and learn about multiple linear regression; ANOVA; logistic regression; random and mixed-effects models; and models for dependent data. Engage in the identification of appropriate statistical methods and interpretation of software output. R programming is introduced. Three lecture hours a week for one semester. Statistics and Data Sciences 324E and 332 may not both be counted. Prerequisite: Statistics and Data Sciences 302F or Statistics and Data Sciences 320E (or Statistics and Data Sciences 302, 304, 306, or 328M).

SDS 325H. Honors Statistics.

An introduction to the fundamental theories, concepts, and methods of statistics. Emphasizes probability models, exploratory data analysis, sampling distributions, confidence intervals, hypothesis testing, correlation and regression, and the use of statistical software. Three lecture hours a week for one semester. Statistics and Data Sciences 325H and Statistics and Scientific Computation 325H may not both be counted. Prerequisite: Admission to the Dean's Scholars Honors Program in the College of Natural Sciences or consent of instructor.

SDS 329C. Practical Linear Algebra I.

Matrix representations and properties of matrices; linear equations, eigenvalue problems and their physical interpretation; and linear least squares and elementary numerical analysis. Emphasis on physical interpretation, practical numerical algorithms, and proofs of fundamental principles. Three lecture hours a week for one semester. Only one of the following may be counted: Mathematics 340L, 341, Statistics and Data Sciences 329C.

SDS 431. Probability and Statistical Inference.

Introduction to probability and statistical inference. Examine events and random experiments; basic rules of probability; joint, conditional, and marginal probability and independence; discrete and continuous random variables; random sampling and estimation; large-sample theory results and central limit theorem-based inferential summaries; and maximum likelihood estimation. Three lecture hours and two laboratory hours a week for one semester. Statistics and Data Sciences 321 and 431 may not both be counted. Prerequisite: Statistics and Data Sciences 315 with a grade of at least C-; and credit with a grade of at least C- or registration for Mathematics 408D, 408L, or 408S.

SDS 334. Intermediate Statistical Methods.

Introduction to applied regression analysis. Explore estimation and inference in multiple regression models; logistic regression; regression for count data; time-to-event models; and case studies in regression modeling in published work, emphasizing both the use and limitations of regression modeling in advancing scientific knowledge. The equivalent of three lecture hours a week for one semester. Prerequisite: Statistics and Data Sciences 431 with a grade of at least C-; Mathematics 340L or 341 or Statistics and Data Sciences 329C with a grade of at least C-; and Computer Science 303E or 312 with a grade of at least C-.

SDS 335. Scientific and Technical Computing.

A comprehensive introduction to computing techniques and methods applicable to many scientific disciplines and technical applications. Covers computer hardware and operating systems, systems software and tools, code development, numerical methods and math libraries, and basic visualization and data analysis tools. Three lecture hours a week for one semester. Statistics and Data Sciences 335 and Statistics and Scientific Computation 335 may not both be counted. Prerequisite: Mathematics 408D or 408M, and prior programming experience.

SDS 336. Practical Machine Learning.

Introduction to machine learning for data science with an emphasis on Python programming. Explore comparing algorithm performance; decision-tree algorithms; classification algorithms; model averaging; unsupervised learning; and neural network. The equivalent of three lecture hours a week for one semester. Prerequisite: Statistics and Data Sciences 334 with a grade of at least C-; and Computer Science 327E with a grade of at least C-.

SDS 150K. Data Analysis Applications.

Introduction to the use of statistical mathematical applications for data analysis. Two lecture hours a week for eight weeks. Statistics and Data Sciences 150K and Statistics and Scientific Computation 150K may not both be counted unless the topics vary. May be repeated for credit when the topics vary. Offered on the pass/fail basis only. Prerequisite: Varies with the topic.

Topic 1: SPSS Software. Statistics and Data Sciences 150K (Topic 1) and Statistics and Scientific Computation 150K (Topic 1) may not both be counted. Offered on the pass/fail basis only. Additional prerequisite: Upper-division standing.
Topic 2: SAS Software. Statistics and Data Sciences 150K (Topic 2) and Statistics and Scientific Computation 150K (Topic 2) may not both be counted. Offered on the pass/fail basis only. Additional prerequisite: Upper-division standing.
Topic 3: Stata Software. Statistics and Data Sciences 150K (Topic 3) and Statistics and Scientific Computation 150K (Topic 3) may not both be counted. Offered on the pass/fail basis only. Additional prerequisite: Upper-division standing.
Topic 4: The R Software Environment. Statistics and Data Sciences 150K (Topic 4) and Statistics and Scientific Computation 150K (Topic 4) may not both be counted. Offered on the pass/fail basis only. Additional prerequisite: Upper-division standing.

SDS 352. Statistical Methods.

Study of simple and multiple regression, fundamentals of experimental design, and analysis of variance methods. May include logistic regression, Poisson regression, resampling methods, introduction to Bayesian methods, and probability models. Includes substantial use of statistical software. Three lecture hours and one laboratory hour a week for one semester. Statistics and Data Sciences 352 and Statistics and Scientific Computation 352 may not both be counted. Prerequisite: One of the following: Mathematics 316, Statistics and Data Sciences 303 (or Statistics and Scientific Computation 303), 304 (or Statistics and Scientific Computation 304), 305 (or Statistics and Scientific Computation 305), or 306 (or Statistics and Scientific Computation 306).

SDS 353. Advanced Multivariate Modeling.

Advanced topics in statistical modeling, including models for categorical and count data; spatial and time-series data; and survival, hazard, and hierarchical models. Extensive use of statistical software to build on knowledge of introductory probability and statistics, as well as multiple regression. Three lecture hours a week for one semester Statistics and Data Sciences 353 and Statistics and Scientific Computation 353 may not both be counted. Prerequisite: Mathematics 408D or 408M; and Statistics and Data Sciences 325H (or Statistics and Scientific Computation 325H) or 352 (Statistics and Scientific Computation 352).

SDS 354. Advanced Statistical Methods.

Explore advanced methods in statistics and data science. Examine modeling data with multilevel (hierarchical) structure and causal inference, including design and analysis strategies. Discuss smoothing

methods; spatial and time series models; additive models; and models for network data. The equivalent of three lecture hours a week for one semester. Prerequisite: Statistics and Data Sciences 334 with a grade of at least C-.

SDS 357. Case Studies in Data Science.

Explore advanced case studies in data science, with an emphasis on the full data analysis pipeline. Examine data collection, identification of data limitations; data privacy; data preparation and exploration; building, using, and evaluating models; creating data products; and communication and persuasion with data. The equivalent of three lecture hours a week for one semester. Prerequisite: Statistics and Data Sciences 334 with a grade of at least C-; credit with a grade of at least C- or registration for Statistics and Data Sciences 336.

SDS 358. Special Topics in Statistics.

Three lecture hours a week for one semester. Statistics and Data Sciences 358 and Statistics and Scientific Computation 358 may not both be counted unless topics vary. May be repeated for credit when the topics vary. Prerequisite: Upper-division standing; additional prerequisites may vary with the topic.

Topic 1: Applied Regression Analysis. Through software application, discussion, and guided instruction, explores simple linear regression - what data is appropriate, how to run analysis, and how to interpret the output. Examines multiple regression with combinations of predictor variables, both continuous and categorical. There will be a discussion/application of ANOVA, prior to preceding on to logistic regression: the prediction of discrete events. Statistics and Data Sciences 358 (Topic: Applied Regression Analysis) and 358 (Topic 1) may not both be counted. Additional prerequisite: One of the following with a grade of at least C-: Statistics and Data Sciences 302, 304, 306, 328M.

SDS 364. Bayesian Statistics.

Introduction to the Bayesian approach for statistical inference. Explore prior, posterior, and predictive distributions: conjugate priors; informative and non-informative priors; models for normal, categorical, and count data; Bayesian computation, including MCMC and the Gibbs sampler; hierarchical models; and Bayesian model checking and model selection. The equivalent of three lecture hours a week for one semester. Prerequisite: Statistics and Data Sciences 431 or 321 with a grade of at least C-; Mathematics 340L or 341 or Statistics and Data Sciences 329C with a grade of at least C-; and credit with a grade of at least C- or registration for Statistics and Data Sciences 334.

SDS 368. Statistical Theory.

Introduction to the mathematical theory of statistics. Explore maximum likelihood estimation, confidence intervals, hypothesis tests and statistical decision theory, tail and concentration bounds, concentration of measure, and nonparametric statistics. The equivalent of three lecture hours a week for one semester. Prerequisite: Statistics and Data Sciences 431 with a grade of at least C-; Mathematics 340L or 341 or Statistics and Data Sciences 329C with a grade of at least C-; and credit with a grade of at least C- or registration for Statistics and Data Sciences 334; and a solid foundation in calculus, probability theory, and linear algebra.

SDS 374C. Parallel Computing for Science and Engineering.

Study of parallel computing principles, architectures, and technologies; and parallel application development, performance, and scalability. Designed to help prepare students to formulate and develop parallel algorithms to implement effective applications for parallel computing systems. Three lecture hours a week for one semester. Statistics and Data Sciences 374C and Statistics and Scientific Computation 374C may not both be counted. Prerequisite: Mathematics 408D or 408M, 340L, and prior programming experience using C or Fortran on Linux or Unix systems.

SDS 374E. Visualization and Data Analysis for Science and Engineering.

Scientific visualization principles, practices, and technologies, including remote and collaborative visualization. Introduces statistical analysis, data mining, and feature detection. Three lecture hours a week for one semester. Statistics and Data Sciences 374E and Statistics and Scientific Computation 374E may not both be counted. Prerequisite: Mathematics 408D or 408M, 340, and prior programming experience using C or Fortran on Linux or Unix systems.

SDS 375. Special Topics in Scientific Computation.

Three lecture hours a week for one semester. Statistics and Data Sciences 375 and Statistics and Scientific Computation 375 may not both be counted unless topics vary. May be repeated for credit when the topics vary. Prerequisite: Upper-division standing; additional prerequisites may vary with the topic.

SDS 378. Introduction to Mathematical Statistics.

Same as Mathematics 378K. Sampling distributions of statistics, estimation of parameters (confidence intervals, method of moments, maximum likelihood, comparison of estimators using mean square error and efficiency, sufficient statistics), hypothesis tests (p-values, power, likelihood ratio tests), and other topics. Three lecture hours a week for one semester. Mathematics 378K and Statistics and Data Sciences 378 may not both be counted. Prerequisite: Mathematics 362K with a grade of at least C-.

SDS 378P. Decision Analytics.

Examine decision theory with utility functions, including the use of probability, optimization, constrained optimization, and linear algebra. Three lecture hours a week for one semester. Only one of the following may be counted: Mathematics 375T (Topic: Decision Analytics), 378P, Statistics and Data Sciences 378P. Prerequisite: Mathematics 362K and Mathematics 378K with a grade of at least C-, or consent of the instructor.

SDS 179R, 279R, 379R, 479R. Undergraduate Research.

Students work on an individual research project under the supervision of one or more faculty members. For each semester hour of credit earned, the equivalent of one lecture hour a week for one semester. May be repeated for credit. Prerequisite: Upper-division standing and consent of instructor.

Graduate Courses

SDS 380C. Statistical Methods I.

Introduction to the fundamental concepts and methods of statistics. Includes descriptive statistics, sampling distributions, confidence intervals, and hypothesis testing. May include simple and multiple linear regression, analysis of variance, and categorical analysis. Use of statistical software is emphasized. Three lecture hours a week for one semester. Statistics and Data Sciences 380C and Statistics and Scientific Computation 380C may not both be counted. Prerequisite: Graduate standing.

SDS 380D. Statistical Methods II.

Continuation of Statistics and Data Sciences 380C (or Statistics and Scientific Computation 380C). Surveys advanced statistical modeling and may include random and mixed effects models, time series analysis, survival analysis, Bayesian methods, and multivariate analysis of variance. Use of statistical software is emphasized. Three lecture hours a week for one semester. Statistics and Data Sciences 380D and Statistics and Scientific Computation 380D may not both be counted. Prerequisite:

Graduate standing, and Statistics and Data Sciences 380C (or Statistics and Scientific Computation 380C) or the equivalent.

SDS 381. Mathematical Methods for Statistical Analysis.

Introduction to mathematical concepts and methods essential for multivariate statistical analysis. Areas may include basic matrix algebra, eigenvalues and eigenvectors, quadratic forms, vector and matrix differentiation, unconstrained optimization, constrained optimization, and applications in multivariate statistical analysis. Three lecture hours a week for one semester. Statistics and Data Sciences 381 and Statistics and Scientific Computation 381 may not both be counted. Prerequisite: Graduate standing and a course in statistics.

SDS 382. Introduction to Probability and Statistics.

Expectation and variance of random variables, conditional probability and independence, sampling distributions, point estimation, confidence intervals, hypothesis tests, and other topics. Three lecture hours a week for one semester. Statistics and Data Sciences 382 and Statistics and Scientific Computation 382 may not both be counted. Prerequisite: Graduate standing and a second-semester calculus class.

SDS 383C. Statistical Modeling I.

An introduction to core applied statistical modeling ideas from a probabilistic, Bayesian perspective. Topics include exploratory data analysis, programming in R, Bayesian probability models, an introduction to the Gibbs sampler, applied regression analysis, and hierarchical models. Three lecture hours a week for one semester. Statistics and Data Sciences 383C and Statistics and Scientific Computation 383C may not both be counted. Prerequisite: Graduate standing.

SDS 383D. Statistical Modeling II.

Use of structured, probabilistic models that incorporate multiple layers of uncertainty to describe real-world systems. Topics include multivariate normal distribution, mixture models, nonparametric Bayesian analysis, advanced hierarchical models and latent-variable models, generalized linear models, and advanced topics in linear and nonlinear regression. Three lecture hours per week for one semester. Statistics and Data Sciences 383D and Statistics and Scientific Computation 383D may not both be counted. Prerequisite: Graduate standing; Economics 392M (Topic 19), Statistics and Data Sciences 384 (or Statistics and Scientific Computation 384), or the equivalent; and 383C (or Statistics and Scientific Computation 383C).

SDS 183K. Data Analysis Applications.

Introduction to the use of statistical or mathematical applications for data analysis. Two lecture hours a week for eight weeks. Statistics and Data Sciences 183K and Statistics and Scientific Computation 183K may not both be counted unless the topics vary. May be repeated for credit when the topics vary. Offered on the credit/no credit basis only. Prerequisite: Graduate standing.

Topic 1: SPSS Software. Statistics and Data Sciences 183K (Topic 1) and Statistics and Scientific Computation 183K (Topic 1) may not both be counted. Offered on the credit/no credit basis only.
Topic 2: SAS Software. Statistics and Data Sciences 183K (Topic 2) and Statistics and Scientific Computation 183K (Topic 2) may not both be counted. Offered on the credit/no credit basis only.
Topic 3: Stata Software. Statistics and Data Sciences 183K (Topic 3) and Statistics and Scientific Computation 183K (Topic 3) may not both be counted. Offered on the credit/no credit basis only.
Topic 4: The R Software Environment. Statistics and Data Sciences 183K (Topic 4) and Statistics and Scientific Computation 183K (Topic 4) may not both be counted. Offered on the credit/no credit basis only.

SDS 384. Topics in Statistics and Probability.

Concepts of probability and mathematical statistics with applications in data analysis and research. Three lecture hours a week for one semester. Statistics and Data Sciences 384 and Statistics and Scientific Computation 384 may not both be counted unless topics vary. May be repeated for credit when the topics vary. Prerequisite: Graduate standing; and Statistics and Data Sciences 382 (or Statistics and Scientific Computation 382), an introductory probability course and a statistics course, or consent of instructor.

Topic 1: Applied Probability. Basic probability theory, combinatorial analysis of random phenomena, conditional probability and independence, parametric families of distributions, expectation, distribution of functions of random variables, and limit theorems. Statistics and Data Sciences 384 (Topic 1) and Statistics and Scientific Computation 384 (Topic 1) may not both be counted.
Topic 2: Mathematical Statistics I. Same as Computational Science, Engineering, and Mathematics 384R and Mathematics 384C. The general theory of mathematical statistics. Includes distributions of functions of random variables, properties of a random sample, principles of data reduction, an overview of hierarchical models, decision theory, Bayesian statistics, and theoretical results relevant to point estimation, interval estimation, and hypothesis testing. Three lecture hours a week for one semester. Only one of the following may be counted: Computational Science, Engineering, and Mathematics 384R, Mathematics 384C, Statistics and Data Sciences 384 (Topic 2). Additional prerequisite: Graduate standing; and Mathematics 362K and 378K, or consent of instructor.
Topic 3: Mathematical Statistics II. Same as Computational Science, Engineering, and Mathematics 384S and Mathematics 384D. Continuation of Computational Science, Engineering, and Mathematics 384R and Mathematics 384C. Three lecture hours a week for one semester. Only one of the following may be counted: Computational Science, Engineering, and Mathematics 384S, Mathematics 384D, Statistics and Data Sciences 384 (Topic 3). Additional prerequisite: Graduate standing; Computational Science, Engineering, and Mathematics 384R, or Mathematics 384C; and Mathematics 362K and 378K, Statistics and Data Sciences 382, or consent of instructor.
Topic 4: Regression Analysis. Same as Computational Science, Engineering, and Mathematics 384T and Mathematics 384G. Simple and multiple linear regression, inference in regression, prediction of new observations, diagnosis and remedial measures, transformations, and model building. Emphasis on both understanding the theory and applying theory to analyze data. Three lecture hours a week for one semester. Only one of the following may be counted: Computational Science, Engineering, and Mathematics 384T, Mathematics 384G, Statistics and Data Sciences 384 (Topic 4). Additional prerequisite: Graduate standing; and Mathematics 362K and 378K, Statistics and Data Sciences 382, or consent of instructor.
Topic 6: Design and Analysis of Experiments. Same as Computational Science, Engineering, and Mathematics 384U and Mathematics 384E. Design and analysis of experiments, including one-way and two-way layouts; components of variance; factorial experiments; balanced incomplete block designs; crossed and nested classifications; fixed, random, and mixed models; and split plot designs. Three lecture hours a week for one semester. Only one of the following may be counted: Computational Science, Engineering, and Mathematics 384U, Mathematics 384E, Statistics and Data Sciences 384 (Topic 6). Additional prerequisite: Graduate standing; and Mathematics 362K and 378K, Statistics and Data Sciences 382, or consent of instructor.
Topic 7: Bayesian Statistical Methods. Fundamentals of Bayesian inference in single-parameter and multi-parameter models for inference and decision making, including simulation of posterior distributions, Markov chain Monte Carlo methods, hierarchical models,

and empirical Bayes models. Statistics and Data Sciences 384 (Topic 7) and Statistics and Scientific Computation 384 (Topic 7) may not both be counted.

Topic 8: Time Series Analysis. Introduction to statistical time series analysis. Includes autoregressive integrated moving average (ARIMA) and more general models, forecasting, spectral analysis, time domain regression, model identification, estimation of parameters, and diagnostic checking. Statistics and Data Sciences 384 (Topic 8) and Statistics and Scientific Computation 384 (Topic 8) may not both be counted. Additional prerequisite: Mathematics 384D.

Topic 9: Computational Statistics. Modern, computation intensive statistical methods, including simulation, optimization methods, Monte Carlo integration, maximum likelihood estimation and expectation-maximization parameter estimation, Markov chain Monte Carlo methods, resampling methods, and nonparametric density estimation. Statistics and Data Sciences 384 (Topic 9) and Statistics and Scientific Computation 384 (Topic 9) may not both be counted.

Topic 10: Stochastic Processes. Concepts and techniques of stochastic processes, with emphasis on the nature of change of variables with respect to time. Includes characterization, structural properties, and inference. Statistics and Data Sciences 384 (Topic 10) and Statistics and Scientific Computation 384 (Topic 10) may not both be counted.

Topic 11: Theoretical Statistics. Examination of asymptotic theory and empirical processes. The former would include minimax theory, Bernstein von Mises theorem, and Bayesian asymptotics. The latter, would include U statistics and robust estimation. Additional prerequisite: Statistics and Data Sciences 384 (Topic 3) or the equivalent; and advanced probability.

SDS 385. Topics in Applied Statistics.

Theories, models, and methods for the analysis of quantitative data. Three lecture hours a week for one semester. Statistics and Data Sciences 385 and Statistics and Scientific Computation 385 may not both be counted unless topics vary. May be repeated for credit when the topics vary. Prerequisite: Graduate standing; and Statistics and Data Sciences 380C (or Statistics and Scientific Computation 380C), 382 (or Statistics and Scientific Computation 382), or consent of instructor.

Topic 1: Experimental Design. Principles, construction, and analysis of experimental designs. Includes one-way classification, randomized blocks, Latin squares, factorial and nested designs, fixed and random effects, multiple comparisons, and analysis of covariance. Statistics and Data Sciences 385 (Topic 1) and Statistics and Scientific Computation 385 (Topic 1) may not both be counted.

Topic 2: Applied Regression. Examination of the foundation of regression analysis. Subjects include simple and multiple linear regression, residual analysis, transformations, building models with real data, testing models, logistic regression, and panel data analysis. Statistics and Data Sciences 385 (Topic 2) and Statistics and Scientific Computation 385 (Topic 2) may not both be counted. Additional prerequisite: Graduate standing, and one of the following: Public Affairs 397, Statistics and Data Sciences 380C (or Statistics and Scientific Computation 380C), 382 (or Statistics and Scientific Computation 382), or consent of instructor.

Topic 3: Applied Multivariate Methods. Introduction to the analysis of multivariate data as applied to examples from the social sciences. Includes multivariate linear models, principal components and factor analysis, discriminant analysis, clustering, and canonical correlation. Statistics and Data Sciences 385 (Topic 3) and Statistics and Scientific Computation 385 (Topic 3) may not both be counted. Additional prerequisite: Statistics and Data Sciences 385 (Topic 2) (or Statistics and Scientific Computation 385 (Topic 2)), or the equivalent.

Topic 4: Analysis of Categorical Data. Methods for analyzing categorical data. Includes categorical explanatory variables within the general linear model, models of association among categorical variables, and models in which the response variable is categorical or is a count. Emphasis on logical similarities across methods. Statistics and Data Sciences 385 (Topic 4) and Statistics and Scientific Computation 385 (Topic 4) may not both be counted.

Topic 5: Structural Equation Modeling. Introduction to the basic concepts, methods, and computing tools used in structural equation modeling. Designed to help students develop a working familiarity with some common statistical procedures and their application through the use of statistical software. Statistics and Data Sciences 385 (Topic 5) and Statistics and Scientific Computation 385 (Topic 5) may not both be counted. Additional prerequisite: Statistics and Data Sciences 385 (Topic 2) (or Statistics and Scientific Computation 385 (Topic 2)), or the equivalent or consent of instructor.

Topic 6: Hierarchical Linear Models. Introduction to multilevel data structures, model building and testing, effect size, fixed and random effects, missing data and model assumptions, hierarchical linear modeling (HLM) logistics, statistical power, and design planning. Statistics and Data Sciences 385 (Topic 6) and Statistics and Scientific Computation 385 (Topic 6) may not both be counted. Additional prerequisite: Statistics and Data Sciences 385 (Topic 2) (or Statistics and Scientific Computation 385 (Topic 2)), or the equivalent or consent of instructor.

Topic 7: Survey Sampling and Methodology. Survey planning, execution, and analysis. Includes the principles of survey research, including sampling and measurement; questionnaire construction and distribution; response effects; validity and reliability; scaling data sources; and data reduction and analysis. Statistics and Data Sciences 385 (Topic 7) and Statistics and Scientific Computation 385 (Topic 7) may not both be counted.

Topic 8: Introduction to Bayesian Methods. A practical introduction to Bayesian statistical interference, with an emphasis on applications in behavioral and measurement research. Examines how Bayesian statistical inference differs from classical inference in the context of simple statistical procedures and models, such as hypothesis testing, analysis of variance (ANOVA), and regression. Statistics and Data Sciences 385 (Topic 8) and Statistics and Scientific Computation 385 (Topic 8) may not both be counted. Additional prerequisite: Statistics and Data Sciences 385 (Topic 2) (or Statistics and Scientific Computation 385 (Topic 2)) or the equivalent, or consent of instructor.

Topic 9: Longitudinal Data Analysis. Applications of models to data collected at successive points in time. Includes latent growth curve models, models for nonlinear growth, discrete-time and continuous-time event history models, multilevel models for change, random coefficient models, and applications of models to event-occurrence data. Statistics and Data Sciences 385 (Topic 9) and Statistics and Scientific Computation 385 (Topic 9) may not both be counted.

Topic 10: Modern Statistical Methods. Introduction to conducting statistical analysis using modern resampling methods, including bootstrapping and Monte Carlo simulation. Emphasis on theoretical understanding and application. Statistics and Data Sciences 385 (Topic 10) and Statistics and Scientific Computation 385 (Topic 10) may not both be counted.

Topic 11: Applied Mathematical Statistics. Introduction to the basic concepts of probability and mathematical statistics. Includes probability distributions and estimation theory and hypothesis testing techniques. Designed for doctoral students who plan to use statistical methods in their research but do not require a highly mathematical investigation of the subject. Statistics and Data Sciences 385 (Topic 11) and Statistics and Scientific Computation 385 (Topic 11) may not both be counted. Additional prerequisite: A calculus course covering integration and differentiation.

Topic 12: Meta-Analysis. Introduction to the statistics used to synthesize results from a set of studies. May include calculation of different effect sizes, calculating pooled estimates using fixed and random effects models, testing moderating variables using fixed and mixed effects models, testing heterogeneity of effect sizes, and assessing and correcting publication bias. Statistics and Data Sciences 385 (Topic 12) and Statistics and Scientific Computation 385 (Topic 12) may not both be counted.
Topic 13: Factor Analysis. Introduction to exploratory and confirmatory factor analysis. May include review of matrix algebra and vector geometry, principal components and principal axis factoring, and factor rotation methods, as well as single-factor and multiple-factor multisample models, multitrait-multimethod technique, and latent means modeling. Emphasis on critiquing current research. Statistics and Data Sciences 385 (Topic 13) and Statistics and Scientific Computation 385 (Topic 13) may not both be counted. Additional prerequisite: Statistics and Data Sciences 385 (Topic 2) (or Statistics and Scientific Computation 385 (Topic 2)), or the equivalent or consent of instructor.
Topic 14: Maximum-Likelihood Statistics. Introduction to the likelihood theory of statistical inference. Includes probability distributions, estimation theory, and applications of maximum-likelihood estimation (MLE) to models with categorical or limited dependent variables, even count models, event history models, models for time-series cross-section data, and models for hierarchical data. Statistics and Data Sciences 385 (Topic 14) and Statistics and Scientific Computation 385 (Topic 14) may not both be counted.
Topic 15: Survival Analysis and Duration Modeling. Focuses on the statistical methods related to the analysis of survival or of time to event data. Emphasis on practical applications in medicine, biology, economics, criminology, sociology, and engineering. May include Kaplan-Meier estimators, semiparametric and parametric regression models, model development, and model adequacy assessment. Statistics and Data Sciences 385 (Topic 15) and Statistics and Scientific Computation 385 (Topic 15) may not both be counted.
Topic 16: Statistical Models for Big Data. An introduction to approaches for modeling large data sets in the biological, social, and physical sciences. Subjects include: linear and generalized-linear models for very large data sets, graphical models, matrix factorization, latent-variable models, large-scale spatial smoothing, online learning, multiple testing, and convex optimization with applications to data analysis. Additional prerequisite: Basic knowledge of linear algebra and multivariable calculus; knowledge of basic probability and statistics at the level of Statistics and Data Sciences 321 and Mathematics 358K; some familiarity with linear regression modeling; programming experience in R, Python, Matlab, or a similar language.

SDS 386C. Probabilistic Graphical Models.

An introduction to statistical learning methods, exploring both the computational and statistical aspects of data analysis. Topics include numerical linear algebra, convex optimization techniques, basics of stochastic simulation, nonparametric methods, kernel methods, graphical models, decision trees, and data resampling. Three lecture hours a week for one semester. Statistics and Scientific Computation 386C and Statistics and Data Sciences 386C may not both be counted. Prerequisite: Graduate standing.

SDS 386D. Monte Carlo Methods in Statistics.

Stochastic simulation for Bayesian inference, designed to develop an understanding of Markov chair Monte Carlo methods and their underlying theoretical framework. Topics include Markov chains, Monte Carlo integration, Gibbs sampler, Metropolis-Hastings algorithms, slice sampling, and sequential Monte Carlo. Three lecture hours a week for one semester. Statistics and Data Sciences 386D and Statistics and Scientific Computation 386D may not both be counted. Prerequisite: Graduate standing; and Economics 392M (Topic 19), Statistics and Data Sciences 384 (or Statistics and Scientific Computation 384), or the equivalent.

SDS 387. Linear Models.

An exploration of practical applications of the projection approach to linear models, building from a review of essential linear algebra concepts to the theory of linear models from a projection-based perspective. Introduction to Bayesian ideas. Additional topics include analysis of variance, generalized linear models, and variable selection techniques. Three lecture hours a week for one semester. Statistics and Data Sciences 387 and Statistics and Scientific Computation 387 may not both be counted. Prerequisite: Graduate standing; Economics 392M (Topic 19: Probability and Statistics), Statistics and Data Sciences 384 (or Statistics and Scientific Computation 384), or the equivalent; and basic coding skills in R, Matlab, or Stata.

SDS 388. Consulting Seminar.

Supervised experience in applying statistical or mathematical methods to real problems. Includes participation in weekly consulting sessions, directed readings in the statistical literature, the ethics of research and consulting, and report writing and presentations. The equivalent of three lecture hours a week for one semester. Statistics and Data Sciences 388 and Statistics and Scientific Computation 388 may not both be counted. May be repeated for credit. Prerequisite: Graduate standing and consent of instructor.

SDS 389. Time Series and Dynamic Models.

Exploration of the general class of state-space models, or dynamic models. Emphasis is placed on the implementation and use of the models presented, with applications focused on the social sciences. Topics include dynamic regression models, the Kalman filter, time series models, multivariate time series models, conditional variance models, Markov chain Monte Carlo algorithms for state-space models, and particle filters. Three lecture hours a week for one semester. Statistics and Data Sciences 389 and Statistics and Scientific Computation 389 may not both be counted. Prerequisite: Graduate standing; Economics 392M (Topic 19), Statistics and Data Sciences 384 (or Statistics and Scientific Computation 384), or the equivalent; and coding skills in R, Matlab, or Stata.

SDS 189R, 289R, 389R, 489R. Graduate Research.

Individual research project supervised by one or more faculty members. For each semester hour of credit earned, the equivalent of one lecture hour a week for one semester. May be repeated for credit. Prerequisite: Graduate standing.

SDS 190. Readings in Statistics.

Faculty directed research seminar. Activities may vary, but will include readings of cutting-edge research papers, discussion of on-going student and faculty projects, and consulting projects. May be repeated for credit. Prerequisite: Graduate standing.

SDS 190C. Biomedical Big Data Seminar.

Faculty directed research seminar and workshop. Activities may vary but will include literature reviews and group discussion, discussion of faculty and student projects, and presentations. One lecture hour a week for one semester. Statistics and Data Sciences 183K (Topic:Biomedical Big Data Seminar) and 190C may not both be counted. May be repeated for credit. Offered on the credit/no credit basis only. Prerequisite: Graduate standing.

SDS 391C. Statistical Machine Learning.

Introduction to statistical machine learning. Subjects include but are not limited to supervised learning (estimation, linear regression,

classification, online and margin based classification approaches), unsupervised learning (K-means, mixture of Gaussians, dimensionality reduction PCA/SVD), and non-parametric methods. Three lecture hours a week for one semester. May be repeated for credit. Prerequisite: Graduate standing; experience in multivariate statistics, probability, linear algebra.

SDS 391D. Data Mining.

Study of various mathematical and statistical aspects of data mining. Includes supervised learning (regression, classification, and support vector machines) and unsupervised learning (clustering, principal components analysis, and dimensionality reduction). Uses technical tools drawn from linear algebra, multivariate statistics, and optimization. Three lecture hours a week for one semester. Statistics and Data Sciences 391D and Statistics and Scientific Computation 391D may not both be counted. Prerequisite: Graduate standing and a linear algebra course.

SDS 292, 392. Introduction to Scientific Programming.

Introduction to programming using both the C and Fortran (95/2003) languages, with applications to basic scientific problems. Covers common data types and structures, control structures, algorithms, performance measurement, and interoperability. For each semester hour of credit earned, one lecture hour a week for one semester. Only one of the following may be counted: Statistics and Data Sciences 222, 322, 292, 392, Statistics and Scientific Computation 222, 322, 292, 392. Prerequisite: Graduate standing and credit or registration for a first-semester calculus course.

SDS 392M. Computational Economics.

Introduction to the development and solution of economic models of growth, macroeconomic fluctuations, environmental economics, financial economics, general equilibrium models, game theory, and industrial economics. Includes neural nets, genetic algorithms and agent-based methods, and stochastic control theory applied to a variety of economic topics. Three lecture hours a week for one semester. Statistics and Data Sciences 392M and Statistics and Scientific Computation 392M may not both be counted. Prerequisite: Graduate standing.

SDS 393C. Numerical Analysis: Linear Algebra.

Same as Computational Science, Engineering, and Mathematics 383C, Computer Science 383C, and Mathematics 383E. Survey of numerical methods in linear algebra: floating-point computation, solution of linear equations, least squares problems, algebraic eigenvalue problems. Three lecture hours a week for one semester. Only one of the following may be counted: Computational Science, Engineering, and Mathematics 383C, Computer Science 383C, Mathematics 383E, Statistics and Data Sciences 393C. Prerequisite: Graduate standing; Computer Science 367 or Mathematics 368K; and Mathematics 340L, 341, or consent of instructor.

SDS 393D. Numerical Analysis: Interpolation, Approximation, Quadrature, and Differential Equations.

Same as Computational Science, Engineering, and Mathematics 383D, Computer Science 383D, and Mathematics 383F. Survey of numerical methods for interpolation, functional approximation, integration, and solution of differential equations. Three lecture hours a week for one semester. Only one of the following may be counted: Computational Science, Engineering, and Mathematics 383D, Computer Science 383D, Mathematics 383F, Statistics and Data Sciences 393D. Prerequisite: Graduate standing; Computational Science, Engineering, and Mathematics 383C, Computer Science 383C, Mathematics 383E, or Statistics and Data Sciences 393C; and Mathematics 427K and 365C, or consent of instructor.

SDS 394. Scientific and Technical Computing.

Comprehensive introduction to computing techniques and methods applicable to many scientific disciplines and technical applications. Covers computer hardware and operating systems, systems software and tools, code development, numerical methods and math libraries, and basic visualization and data analysis tools. Three lecture hours a week for one semester. Statistics and Data Sciences 394 and Statistics and Scientific Computation 394 may not both be counted. Prerequisite: Graduate standing, and a third-semester calculus course or the equivalent; prior programming experience is recommended.

SDS 394C. Parallel Computing for Science and Engineering.

Parallel computing principles, architectures, and technologies. Parallel application development, performance, and scalability. Designed to prepare students to formulate and develop parallel algorithms to implement effective applications for parallel computing systems. Three lecture hours a week for one semester. Statistics and Data Sciences 394C and Statistics and Scientific Computation 394C may not both be counted. Prerequisite: Graduate standing, a third-semester calculus course or the equivalent, a linear algebra course or a matrices course, and prior programming experience using C or Fortran on Linux or Unix systems.

SDS 394D. Distributed and Grid Computing for Science and Engineering.

Distributed and grid computing principles and technologies. Covers common modes of grid computing for scientific applications, developing grid-enabled applications, and future trends in grid computing. Three lecture hours a week for one semester. Statistics and Data Sciences 394D and Statistics and Scientific Computation 394D may not both be counted. Prerequisite: Graduate standing, a third-semester calculus course or the equivalent, a linear algebra course or a matrices course, and prior programming experience using C or Fortran on Linux or Unix systems.

SDS 394E. Visualization and Data Analysis for Science and Engineering.

Scientific visualization principles, practices, and technologies, including remote and collaborative visualization. Introduces statistical analysis, data mining, and feature detection. Three lecture hours a week for one semester. Statistics and Data Sciences 394E and Statistics and Scientific Computation 394E may not both be counted. Prerequisite: Graduate standing, a third-semester calculus course or the equivalent, a linear algebra course or a matrices course, and prior programming experience using C or Fortran on Linux or Unix systems.

SDS 395. Advanced Topics in Scientific Computation.

Three lecture hours a week for one semester. Statistics and Data Sciences 395 and Statistics and Scientific Computation 395 may not both be counted unless topics vary. May be repeated for credit when the topics vary. Prerequisite: Graduate standing; additional prerequisites vary with the topic.

SDS 197R. Statistics Internship.

May be repeated for credit. Offered on the credit/no credit basis only. Prerequisite: Graduate standing and consent of graduate adviser.

SDS 398R. Master's Report.

Preparation of a report to fulfill the requirement for the master's degree under the report option. The equivalent of three lecture hours a week for one semester. Statistics and Data Sciences 398R and Statistics and Scientific Computation 398R may not both be counted. Offered on the credit/no credit basis only. Prerequisite: Graduate standing in statistics

and data sciences, consent of supervising professor, and consent of graduate adviser.

SDS 398T. Supervised Teaching in Statistics and Data Sciences.

Supervised teaching experience; weekly group meetings, individual consultations, and reports. Three lecture hours a week for one semester. Statistics and Data Sciences 398T and Statistics and Scientific Computation 398T may not both be counted. Offered on the credit/no credit basis only. Prerequisite: Graduate standing and appointment as a teaching assistant.

SDS 399W, 699W, 999W. Dissertation.

Statistics and Data Sciences 399W, 699W, 999W and Statistics and Scientific Computation 399W, 699W, 999W may not both be counted. May be repeated for credit. Offered on the credit/no credit basis only. Prerequisite: Admission to candidacy for the doctoral degree.

Professional Courses