STATISTICS AND DATA SCIENCE

Statistics and Data Science is the study of how to collect, analyze, interpret, and communicate data to understand complex phenomena and support decision-making. Statisticians and data scientists develop methods for extracting meaningful insights from data, building models to describe uncertainty, and designing studies and experiments to answer scientific and societal questions.

Undergraduate Degree Programs

The department offers programs leading to a BS degree in Data Science, with coursework emphasizing statistical modeling, computational methods, and data analysis. The curriculum provides a strong foundation for employment in data-driven fields or for graduate study in statistics, data science, and related disciplines.

Options within the program allow students to tailor their studies toward areas such as life sciences, Economics, Environmental Economics and Policy, and Psychological Science. Undergraduate minors in both Statistics and Data Science are available for students majoring in other fields who wish to gain quantitative and analytical skills.

Graduate Degree Programs

Graduate programs leading to the MS and PhD degrees in Statistics are offered, with training in both theoretical and applied areas such as statistical genetics and genomics, spatial and environmental statistics, and statistical learning. A graduate minor in Statistics is offered to complement advanced study in other fields. Students planning to major in statistics at the graduate level should have a minimum of mathematics through multivariable calculus, linear algebra, and an upper-division sequence in mathematical statistics.

Survey Research Center

Established in 1973, the Oregon State University Survey Research Center (http://stat.oregonstate.edu/src/survey-research-center) (OSU-SRC) provides comprehensive survey services including proposal development, questionnaire design and layout, survey administration and data collection, survey analysis and professional report writing. Our staff offers customized options, working with our clients to determine the best approach to collect survey data based on the study objectives, population of interest, and budgetary concerns. Our past and current clients include federal, state, and local agencies, national non-profit organizations, and OSU-affiliated entities. The OSU-SRC maintains several contracts with clients to provide our services on a recurrent basis, from monthly, annually, to ever few years.

Operating as a center for research in survey methodology, the OSU-SRC routinely conducts experiments using self-administered surveys with an aim to contribute to survey methodology research. The OSU-SRC subsequently publishes related material in scientific journals and presents experimental findings at professional meetings. The OSU-SRC provides expertise using survey best practices to maximize response rates and reduce non-response bias. Various sampling plans are examined for each survey to minimize total survey error. The OSU-SRC also offers consulting for OSU community members on research-based survey design and analysis.

Undergraduate Programs

Program Name	SIS Code	Degree Types	Campus Locations
Data Science Major (https://catalog.oregonstate.edu/college-departments/science/statistics-data-science/data-science-bs-hbs/)	A051	BS, HBS	Corvallis, Ecampus
Data Science Major: Advanced Data Science Option (https://catalog.oregonsta college-departments/science/statistics-data-science/data-science-bs-hbs/)	A052	BS, HBS	Corvallis, Ecampus
Data Science Major: Economics Option (https://catalog.oregonstate.edu/college-departments/science/statistics-data-science/data-science-bs-hbs/economics-option/)	A053	BS, HBS	Corvallis, Ecampus
Data Science Major: Environmental Economics and Policy Option (https://catalog.oregonsta college-departments/science/statistics-data-science/data-science-bs-hbs/environmental-economics-policy-option/)	A054	BS, HBS	Corvallis, Ecampus

Data Science Major: Life Science Option (https://catalog.oregonstate.edu/college-departments/science/statistics-data-science/data-science-bs-hbs/life-science-option/)	A055	BS, HBS	Corvallis
Data Science Major: Psychological Science Option (https://catalog.oregonsta college-departments/science/statistics-data-science/data-science-bs-hbs/psychological-science-option/)	A056	BS, HBS	Corvallis
Data Science Minor (https://catalog.oregonstate.edu/college-departments/science/statistics-data-science/data-science-minor/)	A061	Minor	Corvallis, Ecampus
Statistics Minor (https://catalog.oregonsta college-departments/science/statistics-data-science/statistics-minor/)	615	Minor	Corvallis

Graduate Programs

Program Name	SIS Code	Degree Types	Campus Locations
Data Analytics Certificate (https://catalog.oregonstate.edu/college-departments/science/statistics-data-science/data-analytics-graduate-certificate/)	CG17	Certificate	Ecampus
Data Analytics Major (https://catalog.oregonsta college-departments/science/statistics-data-science/data-analytics-ms/)	6160	MS	Ecampus
Statistics Major (https://catalog.oregonstate.edu/college-departments/science/statistics-data-science/statistics-ms-phd/)	6150	MS, PhD	Corvallis
Statistics Minor (https://catalog.oregonsta college-departments/science/statistics-data-science/statistics-minor/)	6150	Minor	Corvallis

Lan Xue, *Interim Department Head*
239 Weniger Hall
Oregon State University
Corvallis, OR 97331-8574
Phone: 541-737-6577
Email: statistics.office@oregonstate.edu
Website: https://stat.oregonstate.edu/

Faculty & Staff

https://stat.oregonstate.edu/people

Data Science (DS)

DS LDEA, LOWER DIVISION ED ABROAD, 1-16 Credits

This course is repeatable for 99 credits.

DS UDEA, UPPER DIVISION ED ABROAD, 1-16 Credits

This course is repeatable for 99 credits.

DS 101, +EXPLORING CAREERS IN DATA SCIENCE, 1 Credit

Explore the Data Science major and the interdisciplinary options within the major. Examine career goals and pathways to data science careers. Explore the concept of professional development. Expose students to data science applications in various fields and prepare them to be critical thinkers of future career paths.
Attributes: CSC1 — Core Ed - Beyond OSU Career Preparation
Prerequisite: SCI 100 (may be taken concurrently) with D- or better or SCI 300 (may be taken concurrently) with D- or better or CORE 100 (may be taken concurrently) with D- or better or CORE 300 (may be taken concurrently) with D- or better or BA 100 (may be taken concurrently) with D- or better or BA 300 (may be taken concurrently) with D- or better or ENGR 110 (may be taken concurrently) with D- or better or ENGR 110H (may be taken concurrently) with D- or better or ENGR 310 (may be taken concurrently) with D- or better or LA 100 (may be taken concurrently) with D- or better or LA 300 (may be taken concurrently) with D- or better or ED 100 (may be taken concurrently) with D- or better or ED 300 (may be taken concurrently) with D- or better or Baccalaureate Core Student with a score of 1
Available via Ecampus

DS 201, +INTRODUCTION TO DATA SCIENCE, 4 Credits

Explores the data life cycle which includes how data are created, collected, imported, described, used and analyzed. Considers how data are used properly and ethically. Examines regression and classification methods, experiments with basic tools in Python, and compares output between regression and classification models. Provides a brief overview of recent data science trends, including its societal impact and limitations.
Attributes: CFQL — Core Ed - Quantitative Literacy & Analysis
Recommended: High School Algebra

DS 231, PYTHON PROGRAMMING FOR DATA SCIENCE, 4 Credits

Focuses on learning Python programming essential for data science. Explores Python's data types and associated methods. Learns to load, clean, merge, transform, and plot data. Develops skills to write Python functions to solve simple problems arising in data science. Gains proficiency with key libraries like Pandas, NumPy, and matplotlib. Adopts best practices for creating reproducible and efficient code.
Prerequisite: DS 201 with D- or better

DS 350, +DATA FOR GOOD, 3 Credits

Analyzes the social impacts of data-driven tools and policies. Examines how systemic bias is reflected in data analyses and data-driven tools. Discusses approaches to scientific reproducibility and to data privacy. Examines ways that data tools often increase structural inequities, and how those effects can be mitigated using either technological or policy approaches.
Attributes: CSDP — Core Ed - Advanced Difference, Power & Oppression
Prerequisite: DS 201 with D- or better

DS 431, STATISTICAL LEARNING FOR DATA SCIENCE, 3 Credits

Introduces supervised and unsupervised machine learning methods, including linear and nonlinear regression methods, model selection techniques, regularization, tree-based methods, support vector machines, and clustering. Situates these methods within a probabilistic statistical framework. Examines how to compare and assess the performance of prediction and classification methods. Provides the tools to apply these methods in the Python programming language and interpret their output.
Prerequisite: DS 231 with D- or better and (ST 412 [D-] or ECON 424 [D-]) and (ST 421 [D-] or ECON 423 [D-])

DS 451, METHODS OF DATA ANALYSIS FOR COMPLEX AND NETWORK DATA, 3 Credits

Focuses on data analysis methods for complex, heterogeneous, and multi-modal data, including network, image, text, and dependent data. Covers building distributed data analysis pipelines, developing neural network models, and applying techniques for network and dependent data analysis. Explores key tools and frameworks in R and Python to design and implement comprehensive analysis pipelines for modern data challenges.
Prerequisite: DS 431 with D- or better

DS 453, BAYESIAN MODELS FOR DATA SCIENCE, 4 Credits

Introduces main concepts of Bayesian analysis from the statistical foundations to model implementation. Reviews and implements a variety of widely used models from a Bayesian perspective including linear regression, Poisson and Negative Binomial regression, and logistic regression. Emphasizes the computational implementation of these models and discusses numerical approximations for posterior inference, including Markov Chain Monte Carlo approaches.
Prerequisite: (ST 421 with D- or better or ECON 423 with D- or better) and (ST 412 [D-] or ECON 424 [D-])

DS 455, CAUSAL INFERENCE FOR EXPERIMENTAL AND OBSERVATIONAL DATA, 3 Credits

Imparts methods to make causal conclusions from data. Builds foundations of causal inference for randomized trials through the perspective of potential outcomes. Extends this foundation to causal inference for observational studies. Develops methodologies for causal inference using regression modeling of real and simulated data for randomized experiments and observational studies.
Prerequisite: (ST 412 with D- or better or ECON 424 with D- or better) and (ST 422 [D-] or ECON 423 [D-])

DS 495, +^CAPSTONE AND CAREER: DATA SCIENCE, 3 Credits

Builds scientific writing skills and synthesizes data science methods from the program into a final data analysis report. Emphasizes communication and presentation skills. Explores writing for different audiences, and producing public-facing portfolios show-casing skills and example data analysis and technical writing. Develops job application materials for career preparation.
Attributes: CSC2 — Core Ed - Beyond OSU Career Engagement; CSWC — Core Ed - Writing Intensive Curriculum (WIC)
Prerequisite: DS 101 with D- or better and DS 431 [D-] and ST 422 [D-] and ST 437 [D-]

DS 553, BAYESIAN MODELS FOR DATA SCIENCE, 4 Credits

Introduces main concepts of Bayesian analysis from the statistical foundations to model implementation. Reviews and implements a variety of widely used models from a Bayesian perspective including linear regression, Poisson and Negative Binomial regression, and logistic regression. Emphasizes the computational implementation of these models and discusses numerical approximations for posterior inference, including Markov Chain Monte Carlo approaches.
Prerequisite: ST 521 with C or better and ST 512 [C]

Statistics (ST)

ST LDEA, LOWER DIVISION ED ABROAD, 0-16 Credits

This course is repeatable for 99 credits.

ST UDEA, UPPER DIVISION ED ABROAD, 0-16 Credits

This course is repeatable for 99 credits.

ST 199, SPECIAL TOPICS, 3 Credits

ST 243Z, +ELEMENTARY STATISTICS I, 4 Credits

A first course in statistics focusing on the interpretation and communication of statistical concepts. Introduces exploratory data analysis, descriptive statistics, sampling methods and distributions, point and interval estimates, hypothesis tests for means and proportions, and elements of probability and correlation. Technology will be used when appropriate.
Attributes: CFQL — Core Ed - Quantitative Literacy & Analysis
Equivalent to: ST 201
Recommended: High school algebra
Available via Ecampus

ST 244, +ELEMENTARY STATISTICS II, 4 Credits

Focuses on the interpretation and communication of statistical concepts. Introduces estimation and hypothesis testing procedures for a variety of categorical and quantitative data scenarios, nonparametric tests, and regression models. Builds on the concepts introduced in ST 243Z. Uses technology when appropriate.
Attributes: CFQL — Core Ed - Quantitative Literacy & Analysis
Prerequisite: ST 201 with D- or better or ST 243Z with D- or better
Equivalent to: ST 202
Available via Ecampus

ST 314, INTRODUCTION TO STATISTICS FOR ENGINEERS, 3 Credits

Probability, common probability distributions, sampling distributions, estimation, hypothesis testing, control charts, regression analysis, experimental design.
Prerequisite: MTH 252Z with D- or better or MTH 252HZ with D- or better or MTH 252 with D- or better or MTH 252H with D- or better
Available via Ecampus

ST 351, INTRODUCTION TO STATISTICAL METHODS, 4 Credits

Examines study designs, descriptive statistics, collecting and recording data, probability distributions, sampling distributions for means and proportions, hypothesis testing and confidence intervals for means and proportions in one- and two-sample inference, and chi-square tests.
Equivalent to: ST 351H
Recommended: High school algebra with statistics; critical thinking, problem solving, and synthesis skills
Available via Ecampus

ST 351H, INTRODUCTION TO STATISTICAL METHODS, 4 Credits

Examines study designs, descriptive statistics, collecting and recording data, probability distributions, sampling distributions for means and proportions, hypothesis testing and confidence intervals for means and proportions in one- and two-sample inference, and chi-square tests.
Attributes: HNRS — Honors Course Designator
Equivalent to: ST 351
Recommended: High school algebra with statistics; critical thinking, problem solving, and synthesis skills

ST 352, INTRODUCTION TO STATISTICAL METHODS, 4 Credits

Randomization tests and other nonparametric tests for one- and two-sample inference, simple and multiple linear regression, correlation, one- and two-way analysis of variance, logistic regression.
Prerequisite: ST 351 with D- or better or ST 351H with D- or better
Available via Ecampus

ST 405, READING AND CONFERENCE, 1-16 Credits

This course is repeatable for 16 credits.

ST 406, PROJECTS, 1-16 Credits

This course is repeatable for 16 credits.

ST 407, SEMINAR, 1 Credit

ST 410, INTERNSHIP, 1-16 Credits

This course is repeatable for 16 credits.

ST 411, METHODS OF DATA ANALYSIS, 4 Credits

Develops practical methods of statistical inference including parametric and nonparametric tools for comparing two samples, one-way analysis of variance, and simple linear regression. Emphasizes non-technical communication of results, including the appropriate scope of inference. Introduces computational tools for manipulating, analyzing, and visualizing data.

Recommended: ST 351

Available via Ecampus

ST 412, METHODS OF DATA ANALYSIS, 4 Credits

Multiple linear regression, including model checking, dummy variables, using regression to fit analysis of variance models, analysis of covariance, variable selection methods.

Prerequisite: ST 411 with D- or better

Recommended: ST 351

ST 413, METHODS OF DATA ANALYSIS, 4 Credits

Principles of experimental design; randomized block and factorial designs; repeated measures; categorical data analysis, including comparison of proportions, tests of homogeneity and independence in cross-classified frequency tables, Mantel-Haenszel test, logistic regression, log-linear regression. Introduction to multivariate statistics.

Prerequisite: ST 412 with D- or better

Recommended: ST 351

ST 415, DESIGN AND ANALYSIS OF PLANNED EXPERIMENTS, 3 Credits

Emphasizes principles of experimental design; uses, construction and analysis of completely randomized, randomized block and Latin square designs; covariates; factorial treatments, split plotting; random effects and variance components.

Prerequisite: ST 352 with D- or better or ST 411 with D- or better or ECON 424 with D- or better

ST 421, INTRODUCTION TO MATHEMATICAL STATISTICS, 4 Credits

Builds the mathematical foundation of probability models, including sample space, events, and random variables. Develops the frameworks for characterizing probability distributions. Performs probability calculations for discrete and continuous random variables, including means, variances, moment generating functions, transformations, conditional, joint, and marginal distributions. Interprets and evaluates the use of common distribution families in real-world settings.

Prerequisite: MTH 252Z with D- or better or MTH 252HZ with D- or better or MTH 252 with D- or better or MTH 252H with D- or better or MTH 228 with D- or better

Available via Ecampus

ST 422, INTRODUCTION TO MATHEMATICAL STATISTICS, 4 Credits

Sampling distributions, Central Limit Theorem, estimation, confidence intervals, properties of estimators, and hypothesis testing.

Prerequisite: ST 421 with D- or better

Recommended: MTH 253Z

Available via Ecampus

ST 431, SAMPLING METHODS, 3 Credits

Estimation of means, totals and proportions; sampling designs including simple random, stratified, cluster, systematic, multistage and double sampling; ratio and regression estimators; sources of errors in surveys; capture-recapture methods.

Recommended: ST 411

Available via Ecampus

ST 436, R PROGRAMMING FOR DATA, 3 Credits

Focus on R programming from a data science perspective. Combine tools from the tidyverse set of packages to import, clean, prepare, and visualize data. Master basic data types, writing functions, automating repetitive tasks, and good practices for producing readable, reusable, and efficient R code.

Recommended: ST 243Z or ST 351, experience working with data in spreadsheets or a point and click interface, or familiarity with calculating summary statistics (e.g., mean, median, standard deviation) and reading basic statistical charts (e.g., barcharts, histograms, scatterplots); critical thinking, problem solving, and reasoning skills

Available via Ecampus

ST 437, DATA VISUALIZATION, 3 Credits

Introduces principles and techniques for effective data visualization. Covers data manipulation, visual encoding, perception, and design principles. Develops skills in critiquing visualizations, proposing improvements, and creating publication-ready visualizations using industry-standard tools. Addresses ethical considerations in data visualization design and communication.

Prerequisite: ST 351 with C or better or ST 314 with C or better or ST 411 with C or better or ECON 424 with C or better

ST 439, SURVEY METHODS, 3 Credits

Discusses survey design, survey administration, and modes of survey data collection. Emphasizes development of the questionnaire including survey question structures, questionnaire flow, and question formulation. Designs survey using Qualtrics given a specific population. Discusses recommended timings of survey delivery and accompanying letters to recruit participants.

Prerequisite: ST 201 with D- or better or ST 243Z with D- or better or ST 351 with D- or better or ST 351H with D- or better or ECON 423 with D- or better

Available via Ecampus

ST 441, PROBABILITY, COMPUTING, AND SIMULATION IN STATISTICS, 4 Credits

Review of probability, including univariate distributions and limit theorems. Random-number generation and simulation of statistical distributions. Bootstrap estimates of standard error. Variance reduction techniques. Emphasis on the use of computation in statistics using the MATLAB programming language.

Prerequisite: ST 422 with D- or better

ST 443, APPLIED STOCHASTIC MODELS, 3 Credits

Development of stochastic models commonly arising in statistics and operations research, such as Poisson processes, birth-and-death processes, discrete-time and continuous-time Markov chains, renewal and Markov renewal processes. Analysis of stochastic models by simulation and other computational techniques.
Prerequisite: ST 421 with D- or better
Recommended: Experience with a high-level programming language or mathematical computation package

ST 499, SPECIAL TOPICS, 1-4 Credits

This course is repeatable for 8 credits.

ST 501, RESEARCH, 1-16 Credits

This course is repeatable for 16 credits.

ST 503, THESIS, 1-16 Credits

This course is repeatable for 999 credits.

ST 505, READING AND CONFERENCE, 1-16 Credits

This course is repeatable for 16 credits.
Available via Ecampus

ST 506, PROJECTS, 1-16 Credits

This course is repeatable for 99 credits.
Available via Ecampus

ST 507, SEMINAR, 1 Credit

This course is repeatable for 99 credits.

ST 509, CONSULTING PRACTICUM, 2 Credits

This course is repeatable for 99 credits.
Recommended: ST 507 and ST 553

ST 510, INTERNSHIP, 1-16 Credits

This course is repeatable for 16 credits.

ST 511, METHODS OF DATA ANALYSIS, 4 Credits

Develops practical methods of statistical inference including parametric and nonparametric tools for comparing two samples, one-way analysis of variance, and simple linear regression. Emphasizes non-technical communication of results, including the appropriate scope of inference. Introduces computational tools for manipulating, analyzing, and visualizing data.
Recommended: ST 351
Available via Ecampus

ST 512, METHODS OF DATA ANALYSIS, 4 Credits

Multiple linear regression, including model checking, dummy variables, using regression to fit analysis of variance models, analysis of covariance, variable selection methods.
Prerequisite: ST 511 with C or better
Recommended: ST 351

ST 513, METHODS OF DATA ANALYSIS, 4 Credits

Principles of experimental design; randomized block and factorial designs; repeated measures; categorical data analysis, including comparison of proportions, tests of homogeneity and independence in cross-classified frequency tables, Mantel-Haenszel test, logistic regression, log-linear regression. Introduction to multivariate statistics.
Prerequisite: ST 512 with C or better
Recommended: ST 351

ST 515, DESIGN AND ANALYSIS OF PLANNED EXPERIMENTS, 3 Credits

Emphasizes principles of experimental design; uses, construction and analysis of completely randomized, randomized block and Latin square designs; covariates; factorial treatments, split plotting; random effects and variance components.
Recommended: ST 352 or ST 411/ST 511 or ECON 424/ECON 524
Available via Ecampus

ST 516, FOUNDATIONS OF DATA ANALYTICS, 4 Credits

Foundations of estimation and hypothesis testing; desirable properties of estimators; maximum likelihood; one- and two-sample problems; theoretical results are explored through simulations and analysis using R.
Recommended: ST 351
Available via Ecampus

ST 517, DATA ANALYTICS I, 4 Credits

Methods for modeling quantitative data and statistical learning--simple and multiple linear regression; linear mixed effects models; data imputation; prediction and cross-validation; scaling up to large datasets. Simulations and data analysis using R.
Prerequisite: ST 516 with C+ or better
Available via Ecampus

ST 518, DATA ANALYTICS II, 4 Credits

Statistical methods and data analysis techniques for count data. Topics include tests for tables of counts, logistic regression, log-linear regression, generalized linear mixed models, and issues for large datasets. Data analysis in R.
Prerequisite: ST 517 with C+ or better
Available via Ecampus

ST 521, INTRODUCTION TO MATHEMATICAL STATISTICS, 4 Credits

Builds the mathematical foundation of probability models, including sample space, events, and random variables. Develops the frameworks for characterizing probability distributions. Performs probability calculations for discrete and continuous random variables, including means, variances, moment generating functions, transformations, conditional, joint, and marginal distributions. Interprets and evaluates the use of common distribution families in real-world settings.
Recommended: Differential and integral calculus
Available via Ecampus

ST 522, INTRODUCTION TO MATHEMATICAL STATISTICS, 4 Credits

Sampling distributions, Central Limit Theorem, estimation, confidence intervals, properties of estimators, and hypothesis testing.
Prerequisite: ST 521 with C or better
Recommended: MTH 253Z
Available via Ecampus

ST 525, APPLIED SURVIVAL ANALYSIS, 3 Credits

Statistical methods for analyzing survival data or time-to-event data, which may be censored and/or truncated. Specific topics can vary term to term, and could include Kaplan-Meier estimator; K-sample hypothesis tests for survival data; Accelerated failure time model; Cox proportional hazard regression model.
Prerequisite: ST 516 with C or better and ST 517 [C] and ST 518 [C]
Available via Ecampus

ST 531, SAMPLING METHODS, 3 Credits

Estimation of means, totals and proportions; sampling designs including simple random, stratified, cluster, systematic, multistage and double sampling; ratio and regression estimators; sources of errors in surveys; capture-recapture methods.
Recommended: ST 411 or ST 511
Available via Ecampus

ST 536, R PROGRAMMING FOR DATA, 3 Credits

Focus on R programming from a data science perspective. Combine tools from the tidyverse set of packages to import, clean, prepare, and visualize data. Master basic data types, writing functions, automating repetitive tasks, and good practices for producing readable, reusable, and efficient R code.
Available via Ecampus

ST 537, DATA VISUALIZATION, 3 Credits

Perceptual principles for displaying data; critique and improvement of data visualizations; use of color in visualization; principles of tidy data; strategies for data exploration; select special topics.
Prerequisite: ST 512 with C or better or ST 517 with C or better or ST 552 with C or better
Recommended: Familiarity with linear regression and using R
Available via Ecampus

ST 538, MODERN STATISTICAL METHODS FOR LARGE AND COMPLEX DATA SETS, 3 Credits

Provides students with the tools and experience to analyze big and messy data and work effectively in a data science team. Covers the tools to handle big data and answer statistical questions based on the data. Includes three big data analysis projects that students work on in groups. Focuses on proper use of modern data analysis techniques related to regression, classification and clustering for data coming from a variety of application fields. R will be the lingua franca.
Prerequisite: ST 512 with C or better or ST 517 with C or better or ST 552 with C or better or ST 412 with C or better
Available via Ecampus

ST 539, SURVEY METHODS, 3 Credits

Discusses survey design, survey administration, and modes of survey data collection. Emphasizes development of the questionnaire including survey question structures, questionnaire flow, and question formulation. Designs survey using Qualtrics given a specific population. Discusses recommended timings of survey delivery and accompanying letters to recruit participants.
Available via Ecampus

ST 541, PROBABILITY, COMPUTING, AND SIMULATION IN STATISTICS, 4 Credits

Review of probability, including univariate distributions and limit theorems. Random-number generation and simulation of statistical distributions. Bootstrap estimates of standard error. Variance reduction techniques. Emphasis on the use of computation in statistics using the S-Plus or MATLAB programming language.
Recommended: ST 422 or ST 522

ST 543, APPLIED STOCHASTIC MODELS, 3 Credits

Development of stochastic models commonly arising in statistics and operations research, such as Poisson processes, birth-and-death processes, discrete-time and continuous-time Markov chains, renewal and Markov renewal processes. Analysis of stochastic models by simulation and other computational techniques.
Recommended: (ST 421 or ST 521) and experience with a high-level programming language or mathematical computation package.

ST 551, STATISTICAL METHODS, 4 Credits

Properties of t, chi-square and F tests; randomized experiments; sampling distributions and standard errors of estimators, delta method, comparison of several groups of measurements; two-way tables of measurements.
Recommended: Concurrent enrollment in MTH 341 and (ST 422 or ST 522)

ST 552, STATISTICAL METHODS, 4 Credits

Simple and multiple linear regression including polynomial regression, indicator variables, weighted regression, and influence statistics, nonlinear regression and linear models for binary data.
Prerequisite: ST 551 with C or better
Recommended: ST 422 or ST 522

ST 553, STATISTICAL METHODS, 4 Credits

Principles and analysis of designed experiments, including factorial experiments, analysis of covariance, random and mixed effect models.
Prerequisite: ST 552 with C or better

ST 555, ADVANCED EXPERIMENTAL DESIGN, 3 Credits

Designs leading to mixed models including split plots, repeated measures, crossovers and incomplete blocks. Introduction to experimental design in industry including confounding, fractional factorials and response surface methodology. Analysis of unbalanced data.
Prerequisite: ST 553 with C or better

ST 557, APPLIED MULTIVARIATE ANALYSIS, 3 Credits

Multivariate data structures, linear combinations; principal components, factor and latent structure analysis, canonical correlations, discriminant analysis; cluster analysis, multidimensional scaling. Not offered every year.
Recommended: (ST 412 or ST 512) and (MTH 252Z or MTH 245)

ST 558, MULTIVARIATE ANALYTICS, 3 Credits

Basics of matrix algebra, principal components analysis, cluster analysis, factor analysis, multidimensional scaling.
Prerequisite: ST 518 with C- or better
Available via Ecampus

ST 559, BAYESIAN STATISTICS, 3 Credits

Bayesian statistics for data analysis. Characterizations of probability; comparative (Bayesian versus frequentist) inference; prior, posterior and predictive distributions; hierarchical modeling. Computational methods include Markov Chain Monte Carlo for posterior simulation.
Recommended: ST 562

ST 561, THEORY OF STATISTICS, 4 Credits

Distributions of functions of random variables, joint and conditional distributions, sampling distributions, convergence concepts, order statistics.
Recommended: ST 422 or ST 522

ST 562, THEORY OF STATISTICS, 4 Credits

Sufficiency, exponential families, location and scale families; point estimation: maximum likelihood, Bayes, and unbiased estimators; asymptotic distributions of maximum likelihood estimators; Taylor series approximations.
Prerequisite: ST 561 with C or better
Recommended: ST 422 or ST 522

ST 563, THEORY OF STATISTICS, 4 Credits

Hypothesis testing: likelihood ratio, Bayesian, and uniformly most powerful tests; similar tests in exponential families; asymptotic distributions of likelihood ratio test statistics; confidence intervals.
Prerequisite: ST 562 with C or better
Recommended: ST 422 or ST 522

ST 565, TIME SERIES, 3 Credits

Analysis of serially correlated data in both time and frequency domains. Autocorrelation and partial autocorrelation functions, autoregressive integrated moving average models, model building, forecasting; filtering, smoothing, spectral analysis, frequency response studies, Offered winter term in even years.
Recommended: (ST 412 or ST 512) and (ST 422 or ST 522)

ST 566, TIME SERIES ANALYTICS, 3 Credits

Focuses on statistical and analytical tools for analyzing data that are observed sequentially over time. Specific topics can vary term to term, and could include methods for exploratory time series analysis, linear time series models (ARMA, ARIMA), forecasting, spectral analysis and state-space models. The focus will be on applied problems, though some mathematical statistics is necessary for a solid understanding of the statistical issues.
Prerequisite: ST 516 with C or better and ST 517 [C] and ST 518 [C]
Available via Ecampus

ST 567, SPATIAL STATISTICS, 3 Credits

Explores the analysis, visualization, and simulation of geostatistical, areal, point-pattern, and spatio-temporal data. Applies computing tools to fit linear, generalized linear, and hierarchical models for inference and prediction. Considers methods to accommodate very large data sets. Examines methods for spatial sampling.
Recommended: MTH 341 and (ST 412 or ST 512) and (ST 422 or ST 522)

ST 591, INTRODUCTION TO QUANTITATIVE GENOMICS, 3 Credits

Provides an overview of how genomic data is generated and analyzed. It focuses on the underlying biological motivation, theoretical concepts, and analytical challenges associated with genomic research, especially the generation of statistics that summarize genomic data. The class is organized as a combination of lectures and group literature review discussions. Students are expected to actively participate in the class. Students from diverse backgrounds, including quantitative, biological, and computations sciences, are encouraged to enroll.
Recommended: ST 411 or ST 511
Available via Ecampus

ST 592, STATISTICAL METHODS FOR GENOMICS RESEARCH, 3 Credits

Lectures include an overview of statistical methods commonly applied in genomics research. Specific methods can vary term to term, and could include cluster analysis, decision trees, dimension reduction tools, regression models, multiple testing adjustment, variable selection methods, etc. Journal clubs include team-based review and presentations of landmark papers in both statistical methodology and genomics research. Research experience includes whole-term collaboration between students from statistics and other disciplines on real projects.
Recommended: ST 411 or ST 511 or a higher level course such as ST 551
Available via Ecampus

ST 595, CAPSTONE PROJECT, 3 Credits

Integrates and applies the analytics skills learned in the MS in Data Analytics program to solve real-world problems and interpret and communicate results. Engages student teams in the entire process of solving data science projects in realistic settings, from placing the problem into appropriate statistical framework to applying suitable analytic methods to the problem. Emphasizes problem solving, written and oral communication skills.
Prerequisite: ST 516 with C or better and ST 517 [C] and ST 518 [C] and ST 558 [C]
This course is repeatable for 6 credits.
Available via Ecampus

ST 599, SPECIAL TOPICS, 1-4 Credits

This course is repeatable for 16 credits.

ST 601, RESEARCH, 1-16 Credits

This course is repeatable for 16 credits.

ST 603, THESIS, 1-16 Credits

This course is repeatable for 999 credits.

ST 606, PROJECTS, 1-16 Credits

This course is repeatable for 16 credits.

ST 623, GENERALIZED REGRESSION MODELS, 3 Credits

Maximum likelihood analysis for frequency data; regression-type models for binomial and Poisson data; iterative weighted least squares and maximum likelihood; analysis of deviance and residuals; over-dispersion and quasilikelihood models; log-linear models for multidimensional contingency tables.
Prerequisite: ST 553 with C or better and ST 563 [C]

ST 625, SURVIVAL ANALYSIS, 3 Credits

Prepares students to understand and analyze survival data. Concepts to be discussed include: hazard function (failure rate function); nonparametric likelihood; empirical processes; empirical distribution function; censoring (mostly right independent censoring); Kaplan-Meier estimator; Bias of the KM estimator; Cox proportional hazards model; Accelerated Failure Time Model; Partial Likelihood; log-rank test.
Prerequisite: ST 553 with C or better or ST 563 with C or better

ST 651, LINEAR MODEL THEORY, 3 Credits

Least squares estimation, best linear unbiased estimation, parameterizations, multivariate normal distributions, distributions of quadratic forms, testing linear hypotheses, simultaneous confidence intervals. Offered alternate years.
Recommended: ST 553 and ST 563

ST 652, LINEAR MODEL THEORY, 3 Credits

Explores advanced topics in linear and generalized linear mixed models: estimation, tests, confidence intervals, prediction, model diagnostics, model selection.
Prerequisite: ST 651 with C or better
Recommended: ST 553 and ST 563

ST 661, ADVANCED THEORY OF STATISTICS, 3 Credits

Exponential families, sufficient statistics; unbiased, equivariant, Bayes, and admissible estimation. Offered alternate years.
Recommended: ST 563 and MTH 511

ST 662, ADVANCED THEORY OF STATISTICS, 3 Credits

Uniformly most powerful, unbiased, similar, and invariant tests. Offered alternate years.
Prerequisite: ST 661 with C or better
Recommended: ST 563 and MTH 511

ST 663, ADVANCED THEORY OF STATISTICS, 3 Credits

First-order and higher-order asymptotics; likelihood ratio, score, and Wald tests; Edgeworth and saddlepoint approximations. Offered alternate years.
Prerequisite: ST 662 with C or better
Recommended: ST 563 and MTH 511