Statistics (ST)

ST 101 Statistics by Example (3 credit hours)
Sampling, experimental design, tables and graphs, relationships among variables, probability, estimation, hypothesis testing. Real life examples from the social, physical and life sciences, the humanities and sports. Credit not allowed if student has prior credit for another ST course

GEP Mathematical Sciences
Typically offered in Fall and Spring

ST 114 Statistical Programming (3 credit hours)
This is an introductory course in computer programming for statisticians using Python. Emphasis is on designing algorithms, problem solving, and forming good coding practices: methodical development of programs from specifications; documentation and style; appropriate use of control structures such as loops, of data types such as arrays; modular program organization; version control. Students will become acquainted with core statistical computational problems through examples and coding assignments, including computation of histograms, boxplots, quantiles, and least squares regression.

Restriction: Freshmen Statistics majors only in Fall; Spring is open to all.
Typically offered in Fall and Spring

ST 295 Special Topics in Statistics (1-3 credit hours)
Presentation of material not normally available in regular course offerings or offering of new courses on a trial basis. Credits and content determined by faculty member in consultation with the Department Head.

Typically offered in Fall, Spring, and Summer

ST 307 Introduction to Statistical Programming- SAS (1 credit hours)
An introduction to using the SAS statistical programming environment. The course will combine lecture and a virtual computing laboratory to teach students how to use the SAS sytem for: basic data input and manipulation; graphical displays of univariate and bivariate data; one- and two-sample analyses of means; simple linear regression; one-way ANOVA. Documentation of code and writing of statistical reports will be included.

Corequisite: ST 305 or ST 312 or ST 372 or Prerequisite: ST 350 or BUS 350
Typically offered in Fall, Spring, and Summer

ST 308 Introduction to Statistical Programming - R (1 credit hours)
Introduction to the statistical programming language R. The course will cover: reading and manipulating data; use of common data structures (vectors, matrices, arrays, lists); basic graphical representations.

Typically offered in Fall and Spring

ST 311 Introduction to Statistics (3 credit hours)
Examining relationships between two variables using graphical techniques, simple linear regression and correlation methods. Producing data using experiment design and sampling. Elementary probability and the basic notions of statistical inference including confidence interval estimation and tests of hypothesis. One and two sample t-tests, one-way analysis of variance, inference for count data and regression. Credit not allowed if student has prior credit for another ST course or BUS 350

GEP Mathematical Sciences
Typically offered in Fall, Spring, and Summer

ST 312 Introduction to Statistics II (3 credit hours)
A further examination of statistics and data analysis. Inference for comparing multiple samples, experimental design, analysis of variance and post-hoc tests. Inference for correlation, simple regression, multiple regression, and curvilinear regression. Analysis of contingency tables and categorical data.

Prerequisite: ST 311 or ST 350 or BUS 350 or ST 370
GEP Mathematical Sciences
Typically offered in Fall and Spring

ST 350/BUS 350 Economics and Business Statistics (3 credit hours)
Introduction to statistics applied to management, accounting, and economic problems. Emphasis on statistical estimation, inference, simple and multiple regression, and analysis of variance. Use of computers to apply statistical methods to problems encountered in management and economics.

Prerequisite: MA 114
Typically offered in Fall, Spring, and Summer

ST 370 Probability and Statistics for Engineers (3 credit hours)
The class is a calculus-based introduction to probability and statistics, with a focus on collection and summary of data, along with making formal inferences and practical conclusions on the basis of data. Topics may include sampling, descriptive statistics, designed experiments, simple and multiple regression, basic probability, discrete and continuous distributions, sampling distributions, hypothesis testing, confidence intervals, one and two-way ANOVA.

Prerequisite: MA 241
Typically offered in Fall, Spring, and Summer

ST 371 Introduction to Probability and Distribution Theory (3 credit hours)
Basic concepts of probability and distribution theory for students in the physical sciences, computer science and engineering. Provides the background necessary to begin study of statistical estimation, inference, regression analysis, and analysis of variance.

Prerequisite: MA 241, Corequisite: MA 242
Typically offered in Fall and Spring

ST 372 Introduction to Statistical Inference and Regression (3 credit hours)
Statistical inference and regression analysis including theory and applications. Point and interval estimation of population parameters. Hypothesis testing including use of t, chi-square and F. Simple linear regression and correlation. Introduction to multiple regression and one-way analysis of variance.

Prerequisite: ST 371
Typically offered in Fall and Spring

ST 401 Experiences in Data Analysis (4 credit hours)
This course will allow students to see many practical aspects of data analysis. Each section of this course will expose students to the process of data analysis in a themed area such as biostatistics or environmental statistics. Students will see problems of data collection and analysis through a combination of classroom demonstrations, hands on computer activities and visits to local industries.

Prerequisite: Permission of Instructor and either ST 311 or ST 305
Typically offered in Summer only

ST 404/BIO 404 Epidemiology and Statistics in Global Public Health (3 credit hours)
This course will provide a general introduction to the quantitative methods used in global health, combining elements of epidemiology and biostatistics. The course will focus on linear and logistic regression, survival analysis, traditional study designs, and modern study designs. Students will learn fundamental principles in epidemiology, including statistical approaches, and apply them to topics in global public health.

Prerequisite: B- or better in ST 311 or ST/BUS 350 or ST 370 or ST 371 and either GPH 201 or GOH 201
Typically offered in Fall only

ST 405/ST 505 Applied Nonparametric Statistics (3 credit hours)
Statistical methods requiring relatively mild assumptions about the form of the population distribution. Classical nonparametric hypothesis testing methods, Spearman and Kendall correlation coefficients, permutation tests, bootstrap methods, and nonparametric regressions will be covered.

Prerequisite: ST 508 or ST 512 or ST 514 or ST 516
Typically offered in Fall only

ST 412/MA 412 Long-Term Actuarial Models (3 credit hours)
Long-term probability models for risk management systems. Theory and applications of compound interest, probability distributions of failure time random variables, present value models of future contingent cash flows, applications to insurance, health care, credit risk, environmental risk, consumer behavior and warranties.

Prerequisite: MA 241 or MA 231, Corequisite: MA 421, BUS(ST) 350, ST 301, ST 305, ST 311, ST 361, ST 370, ST 371, ST 380 or equivalent
Typically offered in Fall only

ST 413/MA 413 Short-Term Actuarial Models (3 credit hours)
Short-term probability models for risk management systems. Frequency distributions, loss distributions, the individual risk model, the collective risk model, stochastic process models of solvency requirements, applications to insurance and businessdecisions.

Prerequisite: MA 241 or MA 231, and one of MA 421, ST 301, ST 305, ST 370, ST 371, ST 380, ST 421.
Typically offered in Summer only

ST 421 Introduction to Mathematical Statistics I (3 credit hours)
First of a two-semester sequence of mathematical statistics, primarily for undergraduate majors in Statistics. Introduction to probability, univariate and multivariate probability distributions and their properties, distributions of functions of random variables, random samples and sampling distributions. Credit is not allowed for both ST 421 and MA 421.

Prerequisite: MA 242
Typically offered in Fall and Spring

ST 422 Introduction to Mathematical Statistics II (3 credit hours)
Second of a two-semester sequence of mathematical statistics, primarily for undergraduate majors in Statistics. Random samples, point and interval estimators and their properties, methods of moments, maximum likelihood, tests of hypotheses, elements of nonparametric statistics and elements of general linear model theory.

Prerequisite: ST 421 or MA 421
Typically offered in Fall and Spring

ST 430 Introduction to Regression Analysis (3 credit hours)
Regression analysis as a flexible statistical problem solving methodology. Matrix review; variable selection; prediction; multicolinearity; model diagnostics; dummy variables; logistic and non-linear regression. Emphasizes use of computer.

Prerequisites: (ST 305 or ST 312 or ST 372) and ST 307 and (MA 303 or MA 305 or MA 405)
Typically offered in Fall and Spring

ST 431 Introduction to Experimental Design (3 credit hours)
Experimental design as a method for organizing analysis procedures. Completely randomized, randomized block, factorial, nested, latin squares, split-plot and incomplete block designs. Response surface and covariance adjustment procedures. Stresses use of computer.

Prerequisite: (ST 305 or ST 312 or ST 372) and ST 307
Typically offered in Fall and Spring

ST 432 Introduction to Survey Sampling (3 credit hours)
Design principles pertaining to planning and execution of a sample survey. Simple random, stratified random, systematic and one- and two-stage cluster sampling designs. Emphasis on statistical considerations in analysis of sample survey data. Class project on design and execution of an actual sample survey.

Prerequisite: (ST 305 or ST 312 or ST 372) and ST 307
Typically offered in Fall and Spring

ST 433/ST 533 Applied Spatial Statistics (3 credit hours)
Introduction to statistical models and methods for analyzing various types of spatially referenced data. The focus is on applications with real data and their analysis with statistical programs such as R and SAS. Students are required to write, modify, and run computer code in order to complete homework assignments and final projects.

P: ST 422 and ST 430
Typically offered in Spring only

ST 434/ST 534 Applied Time Series (3 credit hours)
Statistical models and methods for the analysis of time series data using both time domain and frequency domain approaches. An introduction and review of necessary statistical concepts will be given, and a statistical computing package will be introduced. Analyses of real data sets using statistical software will be emphasized.

Prerequisite: ST 422 and ST 430
Typically offered in Fall only

ST 435/ST 535 Statistical Methods for Quality and Productivity Improvement (3 credit hours)
Use of statistics for quality control and productivity improvement. Control chart calculations and graphing, process control and specification; sampling plans; and reliability. Computer use will be stressed for performing calculations and graphing.

Prerequisite: (ST 305 or ST 312 or ST 372) and ST 307
Typically offered in Spring only

ST 437/ST 537 Applied Multivariate and Longitudinal Data Analysis (3 credit hours)
An introduction to use of statistical methods for analyzing multivariate and longitudinal data collected in experiments and surveys. Topics covered include multivariate analysis of variance, discriminant analysis, principal components analysis, factor analysis, covariance modeling, and mixed effects models such as growth curves and random coefficient models. Emphasis is on use of a computer to perform statistical analysis of multivariate and longitudinal data.

Prerequisite: ST 422 and ST 430
Typically offered in Spring only

ST 440/ST 540 Applied Bayesian Analysis (3 credit hours)
Introduction to Bayesian concepts of statistical inference; Bayesian learning; Markov chain Monte Carlo methods using existing software (SAS and OpenBUGS); linear and hierarchical models; model selection and diagnostics.

Prerequisite: ST 422 and ST 430
Typically offered in Spring only

ST 442/CSC 442 Introduction to Data Science (3 credit hours)
Overview of data structures, data lifecycle, statistical inference. Data management, queries, data cleaning, data wrangling. Classification and prediction methods to include linear regression, logistic regression, k-nearest neighbors, classification and regression trees. Association analysis. Clustering methods. Emphasis on analyzing data, use and development of software tools, and comparing methods.

Prerequisite: (MA 305 or MA 405) and (ST 305 or ST 312 or ST 370 or ST 372 or ST 380) and (CSC 111 or CSC 112 or CSC 113 or CSC 114 or CSC 116 or ST 114 or ST 445)
Typically offered in Fall only

ST 445 Introduction to Statistical Computing and Data Management (3 credit hours)
Detailed discussion of the program data vector and data handling techniques that are required to apply statistical methods. Topics are based on the current content of the Base SAS Certification Exam and typically include: importing, validating, and exporting of data files; manipulating, subsetting, and grouping data; merging and appending data sets; basic detail and summary reporting; and code debugging. Additional topics with practical applications, such as graphics and advanced reporting, may also be introduced. Statistical methods for analyzing data are not covered in this course. Regular access to a computer for homework and class exercises is required. Previous exposure to SAS is expected.

Prerequisite: (ST 305 or ST 312 or ST 372) and ST 307
Typically offered in Fall and Spring

ST 446 Intermediate SAS Programming with Applications (3 credit hours)
This course covers a wide range of SAS skills that build on the topics introduced in ST 445: Introduction to Statistical Computing and Data Management. In particular, many topics related to the Advanced SAS Certifi cation Exam are covered in order to help students prepare for that exam. However, an additional goal of equal importance is to synthesize statistical content such as regression, distributional assumptions for inference, and power from multiple courses through simulation- and graphics-based investigations.

Prerequisite: (MA 305 or MA 405) and (ST 445)
Typically offered in Spring only

ST 451 Sports Analytics (3 credit hours)
Statistical reasoning in sports. Quantitative methods for analyzing observational data collected in professional and amateur sports. Probability concepts, such as joint probability, conditional probability, and Bayes Theorem to model uncertainty in sports. Quantifying variation in estimation for comparative purposes including resampling techniques; statistical significance. Linear and generalized linear regression models. Focus on baseball, basketball, North American football, hockey. Other areas of application include tennis, soccer, and golf.

P: (ST 307 or ST 308 or CSC 111 or CSC 116) and (ST 312 or ST 370 or ST 372)
Typically offered in Fall only

ST 452 Statistical Learning and Data Analytics (3 credit hours)
Basic principles and techniques of data analytics, data wrangling, and visualization. Statistical learning techniques, including (a) supervised learning topics like regression and classification, (b) unsupervised learning techniques like clustering and principal component analysis, (c) an introduction to deep learning, and (d) an Introduction to text mining and natural language processing. The concepts will be implemented with R, and the focus will be on practical case studies with real-world datasets.

Prerequisite: ST 430 and (ST 308 or ST 114 or CSC 111 or CSC 116 or ISE 135)
Typically offered in Fall only

ST 453 Advanced Computing for Statistical Reasoning (3 credit hours)
This course is designed to survey topics and tools needed for an undergraduate statistics major to begin to develop a broad and thorough working knowledge of modern computational techniques for statistical inferences. Statistical methods and the algorithms used to facilitate their computations are motivated by building logical foundations for statistical reasoning. Algorithms surveyed can broadly be categorized as either optimization based or sampling based. Rather than focusing on learning standard software packages for implementing common statistical routines, all codes will be written from scratch using the R programming language (or any other high-level language of the students' choosing). Emphasis is placed on developing a practical understanding of how and why existing methods work, and when to apply a particular method. Some programming proficiency is assumed.

Prerequisite: ST 311 and ST 312 and MA 242 and (MA 305 or MA 405) and ST 308
Typically offered in Spring only

ST 491 Statistics in Practice (3 credit hours)
Mentored experience in applied statistical analysis. Students will work in small groups in collaboration with local scientists to answer real questions about real data. The experience involves mentoring by both the project scientist and the instructor.

P: ST 430
Typically offered in Spring only

ST 495 Special Topics in Statistics (1-6 credit hours)
Offered as needed to present material not normally available in regular departmental course offerings, or for offering new courses on a trial basis.

Typically offered in Fall, Spring, and Summer

ST 497 Professional Experience in Statistics (1-3 credit hours)
Mentored professional experience in statistics. A minimum of 45 hours must be completed for each credit hour earned. The experience must be arranged in advance by the student and approved by the Department of Statistics prior to enrollment. Approval requires completion of the Statistics Department's Experiential Learning Contract, which must be signed by the student, their professional mentor, and their academic advisor. Professional mentors are encouraged to require a research paper or poster presentation as part of the work expectations when appropriate. Students should refer to their curriculum requirements for possible restrictions on the total number of ST 497 credit hours that may be applied to their degree.

Prerequisite: Sophomore Standing. Students are responsible for identifying their own internship mentor and experience.
Typically offered in Fall, Spring, and Summer

ST 498 Independent Study In Statistics (1-6 credit hours)
Detailed investigation of topics of particular interest to advanced undergraduates under faculty direction. Individualized/Independent Study and Research courses require a "Course Agreement for Students Enrolled in Non-Standard Courses" be completed by the student and faculty member prior to registration by the department.

Prerequisite: Six hours of ST
Typically offered in Fall, Spring, and Summer

ST 499 Research Experience in Statistics (1-3 credit hours)
Mentored research experience in statistics. A minimum of 45 hours must be completed for each credit hour earned. The experience must be arranged in advance by the student and approved by the Department of Statistics prior to enrollment. Approval requires completion of the Statistics Department's Experiential Learning Contract, which must be signed by the student, their research mentor, and their academic advisor. Research mentors are encouraged to require a research paper or poster presentation as part of the work expectations when appropriate. Students should refer to their curriculum requirements for possible restrictions on the total number of ST 499 credit hours that may be applied to their degree.

Prerequisite: Sophomore Standing. Students are responsible for identifying their own research mentor and experience.
Typically offered in Fall, Spring, and Summer

ST 501 Fundamentals of Statistical Inference I (3 credit hours)
First of a two-semester sequence in probability and statistics taught at a calculus-based level. Probability: discrete and continuous distributions, expected values, transformations of random variables, sampling distributions. Credit not given for both ST 701 and ST 501. Note: this course will be offered in person (Fall) and online (Summer).

Prerequisite: MA 242 or equivalent
Typically offered in Fall and Summer

ST 502 Fundamentals of Statistical Inference II (3 credit hours)
Second of a two-semester sequence in probability and statistics taught at a calculus-based level. Statistical inference: methods of construction and evaluation of estimators, hypothesis tests, and interval estimators, including maximum likelihood. Credit not given for both ST 702 and ST 502. Note: this course will be offered in person (Spring) and online (Fall).

Prerequisite: ST 501
Typically offered in Fall and Spring

ST 503 Fundamentals of Linear Models and Regression (3 credit hours)
Estimation and testing in full and non-full rank linear models. Normal theory distributional properties. Least squares principle and the Gauss-Markov theorem. Estimability, analysis of variance and co variance in a unified manner. Practical model-building in linear regression including residual analysis, regression diagnostics, and variable selection. Emphasis on use of the computer to apply methods with data sets. Credit not given for both ST 705 and ST 503. Note: this course will be offered in person (Spring) and online (Summer).

P: ST 501 and MA 405 or equivalent (Linear Algebra); C: ST 502
Typically offered in Spring and Summer

ST 505/ST 405 Applied Nonparametric Statistics (3 credit hours)
Statistical methods requiring relatively mild assumptions about the form of the population distribution. Classical nonparametric hypothesis testing methods, Spearman and Kendall correlation coefficients, permutation tests, bootstrap methods, and nonparametric regressions will be covered.

Prerequisite: ST 508 or ST 512 or ST 514 or ST 516
Typically offered in Fall only

ST 507 Statistics For the Behavioral Sciences I (3 credit hours)
A general introduction to the use of descriptive and inferential statistics in behavioral science research. Methods for describing and summarizing data presented, followed by procedures for estimating population parameters and testing hypotheses concerning summarized data.

Prerequisite: Graduate standing
Typically offered in Fall and Spring

ST 511 Statistical Methods For Researchers I (3 credit hours)
Basic concepts of statistical models and use of samples; variation, statistical measures, distributions, tests of significance, analysis of variance and elementary experimental design, regression and correlation, chi-square.

Prerequisite: Graduate Standing
Typically offered in Fall, Spring, and Summer

ST 512 Statistical Methods For Researchers II (3 credit hours)
Covariance, multiple regression, curvilinear regression, concepts of experimental design, factorial experiments, confounded factorials, individual degrees of freedom and split-plot experiments. Computing laboratory addressing computational issues and use of statistical software.

Prerequisite: ST 511 or ST 513 or ST 517
Typically offered in Fall, Spring, and Summer

ST 513 Statistics for Management and Social Sciences I (3 credit hours)
This course introduces important ideas about collecting high quality data and summarizing that data appropriately both numerically and graphically. We explore the use of probability distributions to model data and find probabilities. Estimation of parameters and properties of estimators are discussed. Construction and interpretation of commonly used confidence intervals and hypothesis tests are investigated. Students will gain considerable experience working with data. Software is used throughout the course with the expectation of students being able to produce their own analyses.

Prerequisite: Graduate standing
Typically offered in Fall and Spring

ST 514 Statistics For Management and Social Sciences II (3 credit hours)
This course provides an in-depth study of building, validating, and predicting using regression models. Topics include multiple linear regression models with both continuous and categorical predictors, model selection techniques, and residual diagnostics. Bayesian regression models are also explored. Categorical data analysis is covered including contingency table analysis and logistic regression models. Students will gain considerable experience working with data. Software is used throughout the course with the expectation of students being able to produce their own analyses.

P: ST 511, ST 513, ST 517, or equivalent
Typically offered in Spring and Summer

ST 515 Experimental Statistics for Engineers I (3 credit hours)
An introduction to the foundations of probability theory and mathematical statistics useful for research in engineering. Topics include descriptive statistics, probability, discrete and continuous random variables and probability distributions, joint probability distributions and random samples, point estimation, confidence intervals, hypothesis testing, and analysis of variance.

Prerequisite: Graduate standing
Typically offered in Fall and Spring

ST 516 Experimental Statistics For Engineers II (3 credit hours)
This course is intended to give students a background in the methods of statistical analysis and design of experiments that will assist them in conducting research and analyzing data in engineering. Concentration in this course will be on principles of the design of experiments and analysis of variance and regression including post-hoc tests, inference for simple regression, multiple regression, and curvilinear regression.

Prerequisite: ST 515
Typically offered in Fall and Spring

ST 517 Applied Statistical Methods I (3 credit hours)
Course covers basic methods for summarizing and describing data, accounting for variability in data, and techniques for inference. Topics include basic exploratory data analysis, probability distributions, confidence intervals, hypothesis testing, and regression analysis. This is a calculus-based course. Statistical software is used; however, there is no lab associated with the course. Credit not given for this course and ST 511 or ST 513 or ST 515. This course does NOT count as an elective towards a degree or a minor in Statistics. Note: the course will be offered in person (Fall) and online (Fall and Summer).

Prerequisites: MA 241 or equivalent (Calculus II) and MA 405 or equivalent (Linear Algebra)
Typically offered in Fall and Summer

ST 518 Applied Statistical Methods II (3 credit hours)
This second course in statistics for graduate students is intended to further expand students' background in the statistical methods that will assist them in the analysis of data. Course covers many fundamental analysis methods currently used to analyze a wide array of data, mostly arising from designed experiments. Topics include multiple regression models, factorial effects models, general linear models, mixed effect models, logistic regression analysis, and basic repeated measures analysis. This is a calculus-based course. Statistical software is used, however, there is no lab associated with the course. Credit not given for this course and ST 512 or ST 514 or ST 516. Note: this course will be offered in person (Spring) and online (Fall and Spring).

Prerequisite: ST 517
Typically offered in Fall and Spring

ST 519/EMS 519 Teaching and Learning of Statistical Thinking (3 credit hours)
This course is designed to bridge theory and practice on how students develop understandings of key concepts in data analysis, statistics, and probability. Discussion of students' understandings, teaching strategies and the use of manipulatives and technology tools. Topics include distribution, measures of center and spread, sampling, sampling distribution, randomness, and law of large numbers. Must complete a first level graduate statistics course (ST 507, ST 511, or equivalent) before enrolling.

Prerequisite: ST 507 or ST 511
Typically offered in Spring only
This course is offered alternate even years

ST 520 Statistical Principles of Clinical Trials (3 credit hours)
Statistical methods for design and analysis of clinical trials and epidemiological studies. Phase I, II, and III clinical trials. Principle of Intention-to-Treat, effects of non-compliance, drop-outs. Interim monitoring of clinical trials and data safety monitoring boards. Introduction to meta-analysis. There is also discussion of Epidemiological methods time permitting.

Corequisite: ST 501 or ST 521 or ST 701
Typically offered in Fall only

ST 525 Statistics and Computing for Agricultural Data Science (3 credit hours)
The fundamentals of designed experiments, analysis of variance, and regression modeling. Categorical data analysis including logistic regression will be covered. Regular access to a computer for homework, class exercises, and statistical computing is required. The emphasis in this class is on the practical aspects of statistical modeling. Assignments will concentrate on problem solving rather than formal proofs and derivations.

P: ST 511 or equivalent
Typically offered in Fall and Spring

ST 531 Experimental Design (3 credit hours)
Overview and comparison of observational studies and designed experiments followed by a thorough discussion of design principles. Review of estimation and inference for regression and ANOVA models from an experimental design perspective. Review of design and analysis for completely randomized, randomized complete block, and Latin square designs. Designs and analysis methods for factorial experiments, general blocking structures, incomplete block designs, confounded factorials, split-plot experiments, and fractional factorial designs. Examples used to illustrate application and analysis of these designs.

Prerequisite: ST 512, or ST 515, or ST 516, or ST 517, or ST 703
Typically offered in Fall only

ST 533/ST 433 Applied Spatial Statistics (3 credit hours)
Introduction to statistical models and methods for analyzing various types of spatially referenced data. The focus is on applications with real data and their analysis with statistical programs such as R and SAS. Students are required to write, modify, and run computer code in order to complete homework assignments and final projects.

P: ST 422 and ST 430
Typically offered in Spring only

ST 534/ST 434 Applied Time Series (3 credit hours)
Statistical models and methods for the analysis of time series data using both time domain and frequency domain approaches. An introduction and review of necessary statistical concepts will be given, and a statistical computing package will be introduced. Analyses of real data sets using statistical software will be emphasized.

Prerequisite: ST 422 and ST 430
Typically offered in Fall only

ST 535/ST 435 Statistical Methods for Quality and Productivity Improvement (3 credit hours)
Use of statistics for quality control and productivity improvement. Control chart calculations and graphing, process control and specification; sampling plans; and reliability. Computer use will be stressed for performing calculations and graphing.

Prerequisite: (ST 305 or ST 312 or ST 372) and ST 307
Typically offered in Spring only

ST 537/ST 437 Applied Multivariate and Longitudinal Data Analysis (3 credit hours)
An introduction to use of statistical methods for analyzing multivariate and longitudinal data collected in experiments and surveys. Topics covered include multivariate analysis of variance, discriminant analysis, principal components analysis, factor analysis, covariance modeling, and mixed effects models such as growth curves and random coefficient models. Emphasis is on use of a computer to perform statistical analysis of multivariate and longitudinal data.

Prerequisite: ST 422 and ST 430
Typically offered in Fall and Spring

ST 540/ST 440 Applied Bayesian Analysis (3 credit hours)
Introduction to Bayesian concepts of statistical inference; Bayesian learning; Markov chain Monte Carlo methods using existing software (SAS and OpenBUGS); linear and hierarchical models; model selection and diagnostics.

Prerequisite: ST 422 and ST 430
Typically offered in Spring only

ST 542 Statistical Practice (3 credit hours)
This course will provide a discussion-based introduction to statistical practice geared towards students in the final semester of their Master of Statistics degree. Note: the course will be offered in person (Fall) and online (Spring and Summer).

Prerequisite: (ST 512 or ST 514 or ST 516 or ST 518) and (ST 502 or ST 522 or ST 702)
Typically offered in Fall, Spring, and Summer

ST 544 Applied Categorical Data Analysis (3 credit hours)
This course focuses on the concepts, methods, and models used to analyze categorical data, particularly contingency tables, count data and binary/binomial type of data. The topics covered include Pearson Chi-squared independence test for contingency tables, measures of marginal and conditional associations, small-sample inference, logistic regression models for independent binary/binomial data and many extended models for correlated binary/binomial data including matched data and longitudinal data. The course emphasizes the implementation of methods/models using SAS and the interpretation of the results from the output.

Prerequisite: ST 512 or ST 514 or ST 515 or ST 516
Typically offered in Fall only

ST 546/MA 546 Probability and Stochastic Processes I (3 credit hours)
Mathematical foundations of probability theory. Probabilistic measure theory, random variables and their distributions, construction of expectation. Notions of convergence: almost sure, in probability, in L^p, weak convergence, vague convergence. Conditioning, independence, Borel-Cantelli lemmas, weak and strong laws of large numbers, characteristic functions, central limit theorem, and related concentration inequalities.

Prerequisite: MA 421 and MA 425 or MA 511
Typically offered in Fall only

ST 554 Analysis of Big Data (3 credit hours)
Course discusses current big data management practices and software along with statistical paradigms important for big data and predictive analytics. Literate programming and good programming practices are covered.

Prerequisite: ST 511 or ST 513 or ST 517
Typically offered in Spring only

ST 555 Statistical Programming I (3 credit hours)
An introduction to programming and data management using SAS, the industry standard for statistical practice. Detailed discussion of the program data vector and data handling techniques that are required to apply statistical methods. Topics are based on the current content of the Base SAS Certification Exam and typically include: importing, validating, and exporting of data files; manipulating, subsetting, and grouping data; merging and appending data sets; basic detail and summary reporting; and code debugging. Additional topics with practical applications are also introduced, such as graphics and advanced reporting. Statistical methods for analyzing data are not covered in this course. Regular access to a computer for homework and class exercises is required. Previous exposure to SAS is not expected.

Prerequisite: Graduate standing
Typically offered in Fall, Spring, and Summer

ST 556 Statistical Programming II (3 credit hours)
Statistical procedures for importing/managing complex data structures using SQL, automated analysis using macro programming, basic simulation methods and text parsing/analysis procedures. Students learn SAS, the industry standard for statistical practice. Regular access to a computer for homework and class exercises is required.

P: ST 555 or Base SAS Certification
Typically offered in Spring only

ST 557 Using Technology to Teach and Learn with Data (3 credit hours)
Provide educators with an in-depth introduction to applying technology in teaching with data. Students will explore a variety of technological tools for analyzing and visualizing data, including the role of programming in that process. Students will learn pedagogy to help them structure learning activities using a variety of technologies. Students will learn to identify key design elements in technologies and data visualizations that support pedagogical goals.

P: ST 511, ST 513, ST 517, or equivalent
Typically offered in Fall only

ST 558 Data Science for Statisticians (3 credit hours)
Methods for reading, manipulating, and combining data sources including databases. Custom functions, visualizations, and summaries. Common analyses done by data scientists. Methods for communicating results including dashboards. Regular access to a computer for homework and class exercises is required.

Prerequisites: (ST 511, ST 513, ST 517, or equivalent) and (ST 555 or moderate computer programming experience)
Typically offered in Fall and Summer

ST 561/ECG 561 Applied Econometrics I (3 credit hours)
Introduction and application of econometrics methods for analyzing cross-sectional data in economics, and other social science disciplines, such as OLS, IV regressions, and simultaneous equations models. Students should have had a statistical methods course at the 300 level or above as well as Calculus I and II.

Typically offered in Fall only

ST 562 Data Mining with SAS Enterprise Miner (3 credit hours)
This is a hands-on course using modeling techniques designed mostly for large observational studies. Estimation topics include recursive splitting, ordinary and logistic regression, neural networks, and discriminant analysis. Clustering and association analysis are covered under the topic "unsupervised learning," and the use of training and validation data sets is emphasized. Model evaluation alternatives to statistical significance include lift charts and receiver operating characteristic curves. SAS Enterprise Miner is used in the demonstrations, and some knowledge of basic SAS programming is helpful.

Prerequisite: ST 512 or ST 514 or ST 515 or ST 516 or ST 517
Typically offered in Spring only

ST 563 Introduction to Statistical Learning (3 credit hours)
This course will introduce common statistical learning methods for supervised and unsupervised predictive learning in both the regression and classification settings. Topics covered will include linear and polynomial regression, logistic regression and discriminant analysis, cross-validation and the bootstrap, model selection and regularization methods, splines and generalized additive models, principal components, hierarchical clustering, nearest neighbor, kernel, and tree-based methods, ensemble methods, boosting, and support-vector machines.

Prerequisite: ST 512 or ST 514 or ST 515 or ST 517
Typically offered in Summer only

ST 590 Special Topics (1-6 credit hours)
Typically offered in Fall, Spring, and Summer

ST 630 Independent Study (1-3 credit hours)
Typically offered in Fall, Spring, and Summer

ST 693 Master's Supervised Research (1-9 credit hours)
Instruction in research and research under the mentorship of a member of the Graduate Faculty.

Prerequisite: Master's student
Typically offered in Fall, Spring, and Summer

ST 695 Master's Thesis Research (1-9 credit hours)
Thesis Research

Prerequisite: Master's student
Typically offered in Fall, Spring, and Summer

ST 701 Statistical Theory I (3 credit hours)
Probability tools for statistics: description of discrete and absolutely continuous distributions, expected values, moments, moment generating functions, transformation of random variables, marginal and conditional distributions, independence, orderstatistics, multivariate distributions, concept of random sample, derivation of many sampling distributions.

Typically offered in Fall only

ST 702 Statistical Theory II (3 credit hours)
General framework for statistical inference. Point estimators: biased and unbiased, minimum variance unbiased, least mean square error, maximum likelihood and least squares, asymptotic properties. Interval estimators and tests of hypotheses: confidence intervals, power functions, Neyman-Pearson lemma, likelihood ratio tests, unbiasedness, efficiency and sufficiency.

Prerequisite: ST 701
Typically offered in Spring only

ST 703 Statistical Methods I (3 credit hours)
Introduction of statistical methods. Examples include multiple linear regression, concepts of experimental design, factorial experiments, and random-effects modeling. A computing laboratory addresses computational issues and use of statistical software. This course is a prerequisite for most advanced courses in statistics. This section is restricted to statistics and closely related majors.

R: 17STPHD Students Only
Typically offered in Fall only

ST 704 Statistical Methods II (3 credit hours)
This course will introduce many methods that are commonly used in applications. Examples include: model generation, selection, assessment, and diagnostics in the context of multiple linear regression (including penalized regression); linear mixed models; generalized linear models; generalized linear mixed models; nonparametric regression and smoothing; and finite-population sampling basics. Coverage will include some theory, plus implementation using SAS and/or R.

Prerequisite:ST 703; Corequisites: ST 702 and ST 705
Typically offered in Spring only

ST 705 Linear Models and Variance Components (3 credit hours)
Theory of estimation and testing in full and non-full rank linear models. Normal theory distributional properties. Least squares principle and the Gauss-Markoff theorem. Estimability and properties of best linear unbiased estimators. General linear hypothesis. Application of dummy variable methods to elementary classification models for balanced and unbalanced data. Analysis of covariance. Variance components estimation for balanced data.

Corequisite: ST 702
Typically offered in Spring only

ST 706/MA 706/OR 706 Nonlinear Programming (3 credit hours)
An advanced mathematical treatment of analytical and algorithmic aspects of finite dimensional nonlinear programming. Including an examination of structure and effectiveness of computational methods for unconstrained and constrained minimization. Special attention directed toward current research and recent developments in the field.

Prerequisite: OR(IE,MA) 505 and MA 425
Typically offered in Spring only

ST 715 Theory Of Sampling Applied To Survey Design (3 credit hours)
Principles for interpretation and design of sample surveys. Estimator biases, variances and comparative costs. Simple random sample, cluster sample, ratio estimation, stratification, varying probabilities of selection. Multi-stage, systematic and double sampling. Response errors.

Prerequisite: ST 422, ST 512
Typically offered in Fall only

ST 721/GN 721 Genetic Data Analysis (3 credit hours)
The course aims to provide students with the relevant background knowledge and quantitative skills for conducting genetic data analysis to evaluate the genetic effects of complex traits. The course will focus on statistical methodologies and analytical strategies for population-based association studies with genotype and sequencing data collected from whole genome and exome. The specific topics include genetic variants; genetic identity coefficients and its applications; heritability; Hardy-Weinberg disequilibrium; recombination; linkage disequilibrium and association mapping; genome-wide association studies (GWAS); population substructures; multiple testing; single-variant and multi-variant association methods; next-generation sequencing (NGS) data and rare variant analysis; copy number variant analysis; analysis using summary statistics.

Prerequisite: ST 511 or equivalent
Typically offered in Fall only
This course is offered alternate years

ST 732 Longitudinal Data Analysis (3 credit hours)
Introduction to modeling longitudinal data; Population-averaged vs. subject-specific modeling; Classical repeated measures analysis of variance methods and drawbacks; Review of estimating equations; Population-averaged linear models; Linear mixed effects models; Maximum likelihood, restricted maximum likelihood, and large sample theory; Review of nonlinear and generalized linear regression models; Population-averaged models and generalized estimating equations; Nonlinear and generalized linear mixed effects models; Implications of missing data; Advanced topics (including Bayesian framework, complex nonlinear models, multi-level hierarchical models, relaxing assumptions on random effects in mixed effects models, among others). Implementation in SAS and R.

Prerequisites: ST 702 and ST 705
Typically offered in Spring only

ST 733 Spatial Statistics (3 credit hours)
Introduction to the theory and methods of spatial data analysis including: visualization; Gaussian processes; spectral representation; variograms; kriging; computationally-efficient methods; nonstationary processes; spatiotemporal and multivariate models.

Prerequisite: ST 705
Typically offered in Spring only

ST 740 Bayesian Inference and Analysis (3 credit hours)
Introduction to Bayesian inference; specifying prior distributions; conjugate priors, summarizing posterior information, predictive distributions, hierachical models, asymptotic consistency and asymptotic normality. Markov Chain Monte Carlo (MCMC) methods and the use of exising software(e.g., WinBUGS).

Prerequisite: ST 702
Typically offered in Fall only

ST 745 Analysis of Survival Data (3 credit hours)
Statistical methods for analysis of time-to-event data, with application to situations with data subject to right-censoring and staggered entry, including clinical trials. Survival distribution and hazard rate; Kaplan-Meier estimator for survival distribution and Greenwood's formula; log-rank and weighted long-rank tests; design issues in clinical trials. Regression models, including accelerated failure time and proportional hazards; partial likelihood; diagnostics.

Prerequisite: ST 502 or ST 702
Typically offered in Spring only

ST 746 Introduction To Stochastic Processes (3 credit hours)
Markov chains and Markov processes, Poisson process, birth and death processes, queuing theory, renewal theory, stationary processes, Brownian motion.

Typically offered in Spring only

ST 747/MA 747 Probability and Stochastic Processes II (3 credit hours)
Advanced development of stochastic processes. Conditional expectation, filtrations of sigma-algebras, stopping times. Martingales, associated convergence theorems and inequalities, martingale decomposition, optional stopping. Markov chains including random walks, recurrence versus transience, asymptotic behavior. General Markov processes and the related semigroup operators. Construction and properties of Brownian motion, Donsker's invariance principle. Other potential topics include stationary processes, Birkhoff's ergodic theorem, branching processes, Poisson processes.

Prerequisite: MA(ST) 546
Typically offered in Spring only

ST 748/MA 748 Stochastic Differential Equations (3 credit hours)
Theory of stochastic differential equations driven by Brownian motions. Current techniques in filtering and financial mathematics. Construction and properties of Brownian motion, wiener measure, Ito's integrals, martingale representation theorem, stochastic differential equations and diffusion processes, Girsanov's theorem, relation to partial differential equations, the Feynman-Kac formula.

Prerequisite: MA(ST) 747
Typically offered in Fall only

ST 750/ECG 750 Introduction to Econometric Methods (3 credit hours)
Introduction to principles of estimation of linear regression models, such as ordinary least squares and generalized least squares. Extensions to time series and panel data. Consideration of endogeneity and instrumental variables estimation. Limited dependent variable and sample selection models. Attention to implementation of econometric methods using a statistical package and microeconomic and macroeconomic data sets.

Prerequisite: ST 421; Corequisite: ST 422
Typically offered in Spring only

ST 751/ECG 751 Econometric Methods (3 credit hours)
Introduction to important econometric methods of estimation such as Least Squares, instrumentatl Variables, Maximum Likelihood, and Generalized Method of Moments and their application to the estimation of linear models for cross-sectional ecomomic data. Discussion of important concepts in the asymptotic statistical analysis of vector process with application to the inference procedures based on the aforementioned estimation methods.

Prerequisite: ST 421, ST 422
Typically offered in Fall only

ST 752/ECG 752 Time Series Econometrics (3 credit hours)
The characteristics of macroeconomic and financial time series data. Discussion of stationarity and non-stationarity as they relate to economic time series. Linear models for stationary economic time series: autoregressive moving average (ARMA) models; vector autoregressive (VAR) models. Linear models for nonstationary data: deterministic and stochastic trends; cointegration. Methods for capturing volatility of financial time series such as autoregressive conditional heteroscedasticity (ARCH) models. Generalized Method of Moments estimation of nonlinear dynamic models.

Prerequisite: ECG(ST) 751
Typically offered in Spring only

ST 753/ECG 753 Microeconometrics (3 credit hours)
The characteristics of microeconomic data. Limited dependent variable models for cross-sectional microeconomic data: logit/probit models; tobit models; methods for accounting for sample selection; count data models; duration analysis; non-parametricmethods. Panel data models: balanced and unbalanced panels; fixed and random effects; dynamic panel data models; limited dependent variables and panel data analysis.

Prerequisite: ECG 751
Typically offered in Spring only

ST 756/GN 756 Computational Molecular Evolution (3 credit hours)
Phylogenetic analyses of nucleotide and protein sequence data. Sequence alignment, phylogeny reconstruction and relevant computer software. Prediction of protein secondary structure, database searching, bioinformatics and related topics. Project required.

Prerequisite: GN 311 and ST 511
Typically offered in Fall only
This course is offered alternate years

ST 757/HS 757/GN 757 Quantitative Genetics Theory and Methods (3 credit hours)
The essence of quantitative genetics is to study multiple genes and their relationship to phenotypes. How to study and interpret the relationship between phenotypes and whole genome genotypes in a cohesive framework is the focus of this course. We discuss how to use genomic tools to map quantitative trait loci, how to study epistasis, how to study genetic correlations and genotype-by-environment interactions. We put special emphasis in using genomic data to study and interpret general biological problems, such as adaptation and heterosis. The course is targeted for advanced graduate students interested in using genomic information to study a variety of problems in quantitative genetics.

Prerequisite: ST 511
Typically offered in Fall only
This course is offered alternate even years

ST 758 Computation for Statistical Research (3 credit hours)
Computational tools for research in statistics, including applications of numerical linear algebra, optimization and random number generation, using the statistical language R. A project encompassing a simulation experiment will be required.

Prerequisite: ST 702 and ST 705
Typically offered in Fall only

ST 771/BMA 771/MA 771 Biomathematics I (3 credit hours)
Role of theory construction and model building in development of experimental science. Historical development of mathematical theories and models for growth of one-species populations (logistic and off-shoots), including considerations of age distributions (matrix models, Leslie and Lopez; continuous theory, renewal equation). Some of the more elementary theories on the growth of organisms (von Bertalanffy and others; allometric theories; cultures grown in a chemostat). Mathematical theories oftwo and more species systems (predator-prey, competition, symbosis; leading up to present-day research) and discussion of some similar models for chemical kinetics. Much emphasis on scrutiny of biological concepts as well as of mathematical structureof models in order to uncover both weak and strong points of models discussed. Mathematical treatment of differential equations in models stressing qualitative and graphical aspects, as well as certain aspects of discretization. Difference equation models.

Prerequisite: Advanced calculus, reasonable background in biology
Typically offered in Fall only

ST 772/BMA 772/MA 772 Biomathematics II (3 credit hours)
Continuation of topics of BMA 771. Some more advanced mathematical techniques concerning nonlinear differential equations of types encountered in BMA 771: several concepts of stability, asymptotic directions, Liapunov functions; different time-scales. Comparison of deterministic and stochastic models for several biological problems including birth and death processes. Discussion of various other applications of mathematics to biology, some recent research.

Prerequisite: BMA 771, elementary probability theory
Typically offered in Spring only

ST 773/BMA 773/MA 773/OR 773 Stochastic Modeling (3 credit hours)
Survey of modeling approaches and analysis methods for data from continuous state random processes. Emphasis on differential and difference equations with noisy input. Doob-Meyer decomposition of process into its signal and noise components. Examples from biological and physical sciences, and engineering. Student project.

Prerequisite: BMA 772 or ST (MA) 746
Typically offered in Spring only
This course is offered alternate years

ST 779 Advanced Probability for Statistical Inference (3 credit hours)
Sets and classes, sigma-fields and related structures, probability measures and extensions, random variables, expectation and integration, uniform integrability, inequalities, L_p-spaces, product spaces, independence, zero-one laws, convergence notions, characteristic functions, simplest limit theorems, absolute continuity, conditional expectation and conditional probabilities, martingales.

Prerequisite: ST 702
Typically offered in Spring only

ST 790 Advanced Special Topics (1-6 credit hours)
Typically offered in Fall, Spring, and Summer

ST 793 Advanced Statistical Inference (3 credit hours)
Statistical inference with emphasis on the use of statistical models, construction and use of likelihoods, general estimating equations, and large sample methods. Includes introduction to Bayesian statistics and the jackknife and bootstrap.

Prerequisite: ST 702
Typically offered in Fall only

ST 801 Seminar (1 credit hours)
Typically offered in Fall and Spring

ST 810 Advanced Topics in Statistics (1-3 credit hours)
Typically offered in Fall and Spring

ST 835 Readings (1-3 credit hours)
Typically offered in Spring only

ST 841 Statistical Consulting (1 credit hours)
Participation in regularly scheduled supervised statistical consulting sessions with faculty member and client. Consultant's report written for each session. Regularly scheduled meetings with course instructor and other student consultants to present and discuss consulting experiences.

Prerequisite: ST 512 and ST 702
Typically offered in Spring only

ST 851 Internship in Integrated Biostatistical Training Program for CVD Research (1 credit hours)
This is an external internship component of an NHLBI-funded training grant of the Statistics Department jointly with Duke for which PhD students, who are appointed to the grant, work on research projects at Duke Clinical Research Institute.

Typically offered in Fall only

ST 885 Doctoral Supervised Teaching (1-3 credit hours)
Teaching experience under the mentorship of faculty who assist the student in planing for the teaching assignment, observe and provide feedback to the student during the teaching assignment, and evaluate the student upon completion of the assignment.

Prerequisite: Doctoral student
Typically offered in Fall and Spring

ST 893 Doctoral Supervised Research (1-9 credit hours)
Instruction in research and research under the mentorship of a member of the Graduate Faculty.

Prerequisite: Doctoral student
Typically offered in Fall, Spring, and Summer

ST 895 Doctoral Dissertation Research (1-9 credit hours)
Dissertation Research

Prerequisite: Doctoral student
Typically offered in Fall, Spring, and Summer

ST 896 Summer Dissertation Research (1 credit hours)
For graduate students whose programs of work specify no formal course work during a summer session and who will be devoting full time to thesis research.

Prerequisite: Doctoral student
Typically offered in Summer only

ST 899 Doctoral Dissertation Preparation (1-9 credit hours)
For students who have completed all credit hour requirements, full-time enrollment, preliminary examination, and residency requirements for the doctoral degree, and are writing and defending their dissertations.

Prerequisite: Doctoral student
Typically offered in Fall, Spring, and Summer