APPLIED STATISTICS (APSTA-GE)

APSTA-GE 2001 Statistics for the Social and Behavioral Sciences I (3 Credits)
Typically offered Fall
Introduction to the basic tools of applied statistics, using statistical software as a platform to achieve hands-on experience with real data. Topics include both descriptive (measures of location, variability, correlation, and simple regression) and inferential (probability, central limit theorem, sampling distributions, hypothesis tests, and confidence intervals) statistics. The course is not appropriate for students seeking to learn the mathematical underpinnings of these techniques.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2002 Statistics for Behav and Social Sciences II (2 Credits)
Typically offered Spring
Introduces students to an array of inferential techniques (t-tests, one, and two-way anova, simple and multiple regression, nonparametric methods) using statistical software as a platform to achieve hands-on experience with real data. The course is not appropriate for students seeking to learn the mathematical underpinnings of these techniques.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2003 Interm Quantitative Methods: General Linear Model (3 Credits)
Typically offered Fall
This course is designed to meet the data analytic needs of the doctoral students whose dissertation relies on the analysis of quantitative data. Procedures important to the data analyst are covered including data entry and definition, treating missing data, detecting outliers, and transforming distributions. First term topics include multiple regression, analysis of covariance, repeated measures analysis of variance, and multivariate analysis of variance and covariance. Second term topics emphasize categorical data analysis, odds, rations, standardization, log linear models, logistic regression. Other topics include multinominal logistic models, survival analysis, principle components, and factor analysis. The approach is conceptual with heavy reliance on computer software packages. Appropriate for doctoral students desiring specialized knowledge beyond the introductory sequence.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2004 Introductory Statistical Inference in R (2 Credits)
Typically offered Spring
This course covers regression techniques from a simulation-based perspective, with an emphasis on applications rather than mathematical theory. Topics include linear regression with single and multiple predictors; linear regression assumptions, diagnostics, and interpretation; prediction and inference; transformations and interactions; ANOVA; global tests for coefficients; contingency tables; and information criteria and model comparison. R will be used throughout the course.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2006 Math, Statistics and R Programming Bootcamp (0 Credits)
This course covers material that is required as prerequisite knowledge for many of the advanced courses offered by ASH in their APSTA sequence. Concepts covered include the following: advanced algebra, pre-calculus, basic matrix algebra, and basics in calculus (derivatives and integrals); probability of events, simple probability rules, common probability distributions, expectations; exploratory statistics, correlations, law of large numbers, t-tests, and regression; basic coding in R, RMarkdown conventions, loops, functions, and simple statistical operations.
Grading: Grad Steinhardt Pass/Fail
Repeatable for additional credit: No

APSTA-GE 2011 Supervised and Unsupervised Machine Learning (2 Credits)
Typically offered Spring and January terms
Classification and clustering are important statistical techniques commonly applied in many social and behavioral science research problems. Both seek to understand social phenomena through the identification of naturally occurring homogeneous groupings within a population. Classification techniques are used to sort new observations into preexisting or know groupings while clustering techniques sort the population under study into groupings based on their observed characteristics. Both help to reveal hidden structure that may be used in further analysis. This course will compare and contrast these techniques, including many of their variations, with an emphasis on applications.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2012 Causal Inference (3 Credits)
Typically offered Fall
Course provides students with a basic knowledge of both how to perform analyses and critique the use of some more advanced statistical methods useful in answering policy questions. While randomized experiments will be discussed, the primary focus will be the challenge of answering causal questions using data that do not meet such standards. Several approaches for observational data including propensity score methods, instrumental variables, difference in differences, fixed effects models and regression discontinuity designs will be discussed. Examples from real public policy studies will be used to illustrate key ideas and methods.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2013 Missing Data (2 Credits)
Typically offered Spring
Course provides students with a basic knowledge of missing data analysis, beginning with the types of missing data mechanisms (e.g., missing completely at random). We then discuss the problems with ignoring missing data and examine problems with conventional fixes. Single imputation with noise is contrasted with multiple imputation approaches. Real examples from policy research are given throughout. More advanced topics include pattern mixture models and handling data that are not missing at random
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2014 Stats Analysis of Networks (3 Credits)
Typically offered not typically offered
This course is an introduction to the analysis and modeling of network data. Network analysis is a key tool in understanding relational data - data describing the relationships between pairs and groups of individuals, as well as the global structure of relationships. We will focus on applications to and building tools for research in the social sciences, but the methodology can be extended to other areas. By the end of the course, you should have a working knowledge of basic network analysis tools and be able to use them to analyze your own data.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2015 Applied Spatial Statistics (2 Credits)
Typically offered not typically offered
Spatial data arise when information collected on units that reside in different locations. Common examples include geology, criminology & epidemiology, where the goal may be to identify patterning or clusters (;hot spots') in the outcomes across the terrain being examined. In the social sciences, a similar set of questions & techniques are required, for example in studies of homelessness, poverty, environmental justice, & education. However, spatial data present a novel set of exploratory & modeling challenges, given the unique way in which outcomes are related (correlated) with each other through proximity. This course is an overview of the methods needed to analyze data for which it is suspected that the spatial component plays an important role.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2017 Educational Data Science Practicum (2 Credits)
Typically offered Spring
This intensive laboratory course will focus on doing data analysis projects with real data selected by the students. The core skills are oriented around first framing good research questions, then having these guide interacting with data of all types & varying quality (e.g., web-scraped, or clickstream-based rather than large national surveys) via visualization, principled modeling & evaluation of models using statistical learning techniques such as regression, classification & clustering, & presentation of results, using "reproducible research" tools (e.g., knitr, sweave) in the R programming language.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2018 Advanced Causal Inference Designs and Applications (3 Credits)
This course builds statistical and analytic skills in advanced causal inference techniques, focusing on advanced topics and applications in regression discontinuity designs and in matching estimators. We go in depth with the theory, assumptions and plausibility checks, applications, recent advances, combinations of these designs with other techniques, and complications of the designs. Students learn advanced techniques hands-on, and produce a journal-quality manuscript by the end of the semester.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2040 Multi-Level Modeling Growth Curve (2 Credits)
Typically offered Fall
This is a course on models for multi-level growth curve data. These data arise in longitudinal designs, which are quite common to education and applied social, behavioral and policy science. Traditional methods, such as OLS regression, are not appropriate in this settings, as they fail to model the complex correlational structure that is induced by these designs. Proper inference requires that we include aspects of the design in the model itself. Moreover, these more sophisticated techniques allow the researcher to learn new and important characteristics of the social and behavioral processes under study. In this module, we will develop and fit a set of models for longitudinal designs (these are often called growth curve models). The course assignments will use state of the art statistical software to explore, fit and interpret the models.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2042 Multi-Level Modeling: Nested Data/Longitudinal Data (2 Credits)
Typically offered Spring
This is a course on models for multi-level nested data. These data arise in nested designs, which are quite common to education and applied social, behavioral and policy science. Traditional methods, such as OSL regression, are not appropriate in this setting, as they fail to model the complex correlational structure that is induced by these designs. Proper inference requires that we include aspects of the design in the model itself. Moreover, these more sophisticated techniques allow the researcher to learn new and important characteristics of the social and behavioral processes under study. In this module, we will develop and fit a set of models for nested designs (these are sometimes called hierarchical linear models). The course assignments will use state of the art statistical software to explore, fit and interpret the models.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2044 Generalized Linear Models and Extensions (2 Credits)
Typically offered Spring
A second year course in advanced statistical techniques that covers useful quantitative tools in health & policy research. Assuming a strong foundation in regression & the general linear model, this course focuses on data analysis that utilizes models for categorical, discrete or limited outcomes that are commonly seen in health & policy studies. Examples include health status, number of clinic visits, etc. In this course students will also learn the principles of likelihood-based inference, which will assist them in some of the more advanced statistics courses.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2047 Messy Data and Machine Learning (3 Credits)
Typically offered not typically offered
This course is designed to expose students to the complex real-world datasets commonly used in machine learning applications. The course provides an accessible introduction to supervised machine learning, while covering aspects of data collection and cleaning. Specific topics include model construction, evaluation, and regularization, as well as web scraping, text data, feature construction, and measurement error. Students complete short assignments, longer homework sets, and a final project.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2062 Ethics of Data Science (3 Credits)
Typically offered Spring
Course is designed to build students' ethical imaginations and skills for collecting, storing, sharing and analyzing data derived from human subjects including data used in algorithms. The course provides historical background to understand the tenets of informed consent, discrimination, and privacy. Using case study design, students will explore current applications of quantitative reasoning in organizations, algorithmic transparency, and unintended automation of discrimination via data that contains biases rooted in race, gender, class, and other characteristics. cc: Dr.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2085 Basic Statistics (3 Credits)
Typically offered Fall and Spring
This introductory course prepares students to use statistics for data analysis. The course makes use of statistical computer software to achieve hands-on experience with data. This course covers methods for displaying and describing data as well as statistical inference. Topics include frequency distributions and their graphical representations, percentiles, measures of central tendency and dispersion, hypothesis tests, analysis of variance, correlation, and simple regression.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2086 Basic Statistics II (3 Credits)
Typically offered Fall
The second semester builds on the foundation of the first and covers particular methods of statistical inference that rely on the normal t, F, and chi-square distributions to test hypotheses about means, variances, correlations, and proportions.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2093 Psychometric theory and applications (3 Credits)
Typically offered Spring
Reviews and expands on the topics of measurement and reliability for psychological and educational test data. Begins with classical test theory, moves onto unidimensional and multidimensional factor models for continuous data, then item response theory for dichotomous and ordered categorical data. Well-suited for students who have collected test/questionnaire data and want to analyze measurement properties of the test (e.g., reliability, dimensionality) and obtain summary scores for each respondent that can be used for reporting or as variables in further analyses.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2094 Modern Approaches in Measurement (3 Credits)
Typically offered Spring
The course provides students with the software skills and theoretical knowledge required to apply latent construct measurement techniques in R. The course assumes a working knowledge of R, linear regression, and basic probability. Continuous and categorical latent variables, dimensionality reduction, clustering, and finite mixture and diagnostic classification models are covered. The course focuses on cross-sectional applications of measurement techniques, with examples drawn from education, psychology, social science.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2110 Large Databases in Applied Research (3-4 Credits)
Typically offered Spring
This course is designed to serve as a bridge between introductory statistics/econometrics and practical work with real, large-scale databases. The analytic skills taught in this course are broadly applicable to research in education, education policy, and the social, behavioral, and health sciences. Emphasis throughout the course is on hands-on data preparation, workflow, and modeling using the Stata statistical software package.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2122 Frequentist Inference (2 Credits)
Typically offered Spring
This is a course in the intermediate and advanced foundations of statistical inference in the context of applied research. Assuming some prior exposure to probability and statistics, this course will first cover topics such as the principles of estimation and hypothesis testing, and the general and generalized linear models, including scientific computation. This course thoroughly explores the frequentist approach to inference. The student will be expected to understand the mathematical theory, implement related statistical algorithms in statistical programming language such as R, and interpret models and parameters in the context of applied statistical analysis of real data.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2123 Bayesian Inference (2 Credits)
Typically offered Spring
This is a course in intermediate and advanced foundations of statistical inference in the context of applied research and covers Bayesian workflow, conjugate models, MCMC, prediction and model evaluation, Bayesian estimation of GLMs, and introduction to hierarchical/multilevel models. Through this course, students gain an understanding of mathematical theory, implement related statistical algorithms in R and Stan, and interpret models and parameters in the context of applied statistical analysis of real data. Prior exposure to probability and statistics required.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2134 Experimental & Quasi Experimental Design (3 Credits)
Typically offered not typically offered
Application of basic and more complex experimental designs such as block designs, factorial, latin square, and repeated measures as well as single case designs. Students will learn the basis for these designs, power and sample size calculations, and related approaches to statistical analysis.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2135 Data-Driven Methods for Policy Evaluation (3 Credits)
Data-centric technologies are transforming public policy, enabling new approaches for evaluating policies both retrospectively and prospectively; for detecting discriminatory practices; and for auditing and designing "fair" algorithmic systems. While current computational and statistical methods often promise increased efficiency, equity, and transparency, their use also raises complex legal, social, and ethical questions. In this course, we will discuss such methods in a variety of applications, and will examine the relationships between law, policy, and data.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2139 Survey Research Methods (3 Credits)
Typically offered Fall
The survey is the leading mechanism for collecting information on a wide array of topics in our data-driven world. This course is designed to introduce students to the fundamental aspects of the survey & ways for evaluating this form of data collection. Principal topics include: survey design; coverage, sampling, & non-response; modes of data collection; questionnaire construction & evaluation. Throughout this course, students will be given opportunities to engage in actual survey research activities.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2310 Internship (2 Credits)
Typically offered Fall, Spring, and Summer terms
In the internship, students will gain experience working with "real world" data, working with an approved faculty member, local firm or organization. Students will receive practical training focused on the kinds of issues that researchers face in collecting & analyzing data. This course will not only enhance the tools & techniques students develop, but will also possibly lead to employment opportunities after graduation.
Grading: Grad Steinhardt Pass/Fail
Repeatable for additional credit: Yes

APSTA-GE 2331 Data Science for Social Impact (3 Credits)
Typically offered Fall
Course focuses on the competencies required and the issues that arise Course focuses on how analysts use data and quantitative evidence to impact policy and practice. Students will learn how to gather and analyze data to address questions about program efficacy and efficient targeting of resources. Topics will include how to choose organizational partners, implement change, build trust with organizations and civic agencies, satisfy the needs of stakeholders and manage legal, ethical, and logistical constraints. Students will discuss real case studies and appropriate ways to address them.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2351 Practicum in Applied Statistics: Applied Probability (3 Credits)
Typically offered Fall
This course will first cover Kolgomorov's axioms of probabilities, basics of set theory, discrete combinatorial probability, Bayes' theorem, probability distributions and their properties and assumptions of dependence and independence, followed by the foundational topics of statistics: sampling distributions, the law of large numbers and the central limit theorem. This course will mix theoretical approaches with simulation-based illustrations of these main topics. The student will solve via analytical and simulation based approaches in statistical programming language R.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2352 Practicum in Applied Statistics: Statistical Computing (1-3 Credits)
Typically offered Fall
This course will introduce the student to statistical programming and simulation using R. Students will first understand variables, data structures, program flow (e.g., conditional execution, looping) and functional programming, then apply these skills to answer interesting statistical questions involving the comparison of groups. Most statistical analysis will be motivated via simulations, rather than mathematical theory. The course content (programming and data analysis) requires significant outside reading and programming.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2353 Practicum in Applied Statistics: SocSci Research Methodology (1-2 Credits)
Typically offered not typically offered
This course will introduce the student to active research in the social sciences, emphasizing the connections between the substantive research & appropriate statistical methods. To fulfill the course requirements, students will attend research seminars offered through PRIISM & other departments/centers & engage in discussions in a weekly class. Concurrently, students will gain familiarity with academic papers that use quantitative methodology.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2354 Applied Data Analytics for Public Policy (3 Credits)
Typically offered Spring
The goal of the Applied Data Analytics class is to develop the key data analytics skill sets necessary to harness the wealth of newly-available data. Its design offers hands-on training in the context of real microdata. The main learning objectives are to apply new techniques to analyze social problems using and combining large quantities of heterogeneous data from a variety of different sources. It is designed for graduate students who are seeking a stronger foundation in data analytics.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2355 Data Science Translation: Writing and Visualization (3 Credits)
Typically offered occasionally
The goal of this course is to learn how to effectively, honestly, and persuasively communicate about empirical research. Students develop competencies in writing, visualization, and oral presentation of technical material to both technical and lay audiences. Students learn practical strategies with ample opportunity to practice new skills. Assignments and discussion emphasize and explore the tension between concision and accuracy. Students receive feedback from the instructor, peers, and exemplars of intended audiences (e.g. employers who hire data scientists).
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2357 Practicum in Applied Statistics: Social Science Research Methods (1 Credit)
This course introduces students to active research in the social sciences, emphasizing the connections between the substantive research and appropriate statistical methods. Students attend research seminars offered through PRIISM and other departments/centers and engage in weekly discussions. Students gain familiarity with academic papers that use quantitative methodology.
Grading: Ugrd Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2358 Practicum in Applied Statistics: Interactive Data Sci using R Shiny (3 Credits)
This lab course focuses on the practical elements of interactive tool development for data analysis, with specific emphasis on Shiny, an R package for building interactive web apps, including dashboards, interactive visualizations, learning modules, and simulation tools. Students learn coding skills, principles and processes of user centered design, and develop their own portfolio of increasingly complex applications. Special emphasis on communicating results from statistical models and on tools designed expressly for learning. Prerequisite: some experience with R.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 2401 Statistical Consulting Research Seminar (3 Credits)
Typically offered Fall
This course is designed to assist graduate students in the quantitative methods specific to the design and analysis of their theses. In this seminar format, under the guidance of one or more statistical faculty members, students will have opportunity to present and defend their scholarly work-in-progress. They will also be required to critique and provide constructive suggestions for their fellow students. The focus of critiques will be on the research methodology and other statistical issues. Students will additionally benefit from being able to observe how the participating faculty diagnosis and solve statistical issues that arise in others' presented work and to benefit from this advice in their own work. In essence this course provides training in statistical consulting along with detailed feedback on one's dissertation research
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No

APSTA-GE 3208 Management and Ethics of Data (3 Credits)
This course provides an introduction to critical and ethical issues surrounding data. As education leaders, student data is used in various ways for multiple purposes. Recognizing the person behind the numbers, this course covers topics including (1) Research and applied ethics, (2) Concepts of privacy and publicity, (3) Issues in data collection and data mining and (4) Lifecycle of data.
Grading: Grad Steinhardt Graded
Repeatable for additional credit: No