Recommended Courses for the Analytics-Focused Student

BEPP 280/780
Applied Data Analysis

This course will examine how and when data can be used specifically to infer whether there is a causal relationship between two variables. We will emphasize (a) the critical role of an underlying economic theory of behavior in interpreting data and guiding analysis, as well as (b) a range of advanced techniques for inferring causality from non-experimental data, such as regression discontinuity designs, matching estimators, instrumental variables, synthetic controls, heterogeneity modeling and natural experiments. The issue of causality, and the relevance of thinking about models and methods for inferring causality, is just as central and important for "Big Data" as it is when working with traditional data sets in business and public policy. Though each lecture will contain some proofs and derivations (only calculus will be required), the emphasis will be on understanding the underlying concepts, the practical use, implications and limitations of techniques. Students will work intensively with data, drawing from examples in business and public policy. All analysis will be conducted using STATA or R. The goals of the course are for students to become expert consumers able to interpret and evaluate empirical studies as well as expert producers of convincing empirical analysis themselves. Course is well suited for students

Taught by: Robert Jensen.

*Approved course in the Business Analytics MBA major

BSTA 630
Statistical Methods and Data Analysis I

This first course in statistical methods for data analysis is aimed at first-year Biostatistics students. It focuses on the analysis of continuous data. Topics include descriptive statistics (measures of central tendency and dispersion, shapes of distributions, graphical representations of distributions, transformations, and testing for goodness of fit); populations and sampling (hypotheses of differences and equivalence, statistical errors); one- and two-sample t tests; analysis of variance; correlation; nonparametric tests on means and correlations; estimation (confidence intervals and robust methods); categorical data analysis (proportions; statistics and test for comparing proportions; test for matched samples; study design); and regression modeling (simple linear regression, multiple regression, model fitting and testing, partial correlation, residuals, multicollinearity). Examples of medical and biologic data will be used throughout the course, and use of computer software demonstrated.

Prerequisite(s): Multivariable calculus and linear algebra, BSTA 620 (may be taken concurrently) and permission of instructor.

CIS 190
C++ Programming

This course will provide an introduction to programming in C++ and is intended for students who already have some exposure to programming in another language such as Java, C++ provides the programmer with a greater level of control over machine resources and are commonly used in situations where low level access or performance are important. This course will illuminate the issues associated with programming at this level and will cover issues such as explicit memory management, pointers, the compilation process and debugging. The course will involve several programming projects which will provide students with the experience they need to program effectively in these languages. This course assumes programming experience equivalent to CIS 110, CIS 120 or ESE 112.

Prerequisite: CIS 240

CIS 191
Using and Understanding Unix and Linux

Unix, in its many forms, runs much of the world's computer infrastructure, from cable modems and cell phones to the giant clusters that power Google and Amazon. This half-credit course provides a thorough introduction to Unix and Linux. Topics will range from critical basic skills such as examining and editing files, compiling programs and writing shell scripts, to higher level topics such as the architecture of Unix and its programming model. The material learned is applicable to many classes, including CIS 240, CIS 331, CIS 341, CIS 371, and CIS 380.

Prerequisite: CIS 110 or equivalent

CIS 192
Python Programming

Python is an elegant, concise, and powerful language that is useful for tasks large and small. Python has quickly become a popular language for getting things done efficiently in many in all domains: scripting, systems programming, research tools, and web development. This course will provide an introduction to this modern high-level language using hands-on experience through programming assignments and a collaborative final application development project.

Prerequisite: CIS 120 or ESE 112

These are .5 credit courses on programming languages offered by the Engineering school. More information here.

CIS 261
Discrete Probability, Stochastic Processes, and Statistical Inference

The purpose of this course is to provide a 1 CU educational experience which tightly integrates the theory and applications of discrete probability, discrete stochastic processes, and discrete statistical inference in the study of computer science. The intended audience for this class is both those students who are CS majors as well as those intending to be CS majors. Specifically, it will be assumed that the students will know: Set Theory, Mathematical Induction, Number Theory, Functions, Equivalence Relations, Partial-Order Relations, Combinatorics, and Graph Theory at the level currently covered in CIS 160. This course could be taken immediately following CIS 160. Computation and Programming will play an essential role in this course. The students will be expected to use the Maple programming environment in homework exercises which will include: numerical and symbolic computations, simulations, and graphical displays.
Course usually offered in fall term

Prerequisite: CIS 160

CIS 419/519
Introduction to Machine Learning

Machine learning has been essential to the success of many recent technologies, including autonomous vehicles, search engines, genomics, automated medical diagnosis, image recognition, and social network analysis, among many others. This course will introduce the fundamental concepts and algorithms that enable computers to learn from experience, with an emphasis on their practical application to real problems. This course will introduce supervised learning (decision trees, logistic regression, support vector machines, Bayesian methods, neural networks and deep learning), unsupervised learning (clustering, dimensionality reduction), and reinforcement learning. Additionally, the course will discuss evaluation methodology and recent applications of machine learning, including large scale learning for big data and network analysis.

CIS 450/550
Database & Information Systems

This course provides an introduction to the broad field of database and information systems, covering a variety of topics relating to structured data, ranging from data modeling to logical foundations and popular languages, to system implementations. We will study the theory of relational and XML data design; the basics of query languages; efficient storage of data, execution of queries and query optimization; transactions and updates; web-database development; and "big data" and NoSQL systems. The course assumes mathematical and programming experience equivalent to CIS160 and CIS121.

Prerequisites: CIS 121, CIS 160

CIS 521
Artificial Intelligence

This course investigates algorithms to implement resource-limited knowledge-based agents which sense and act in the world. Topics include, search, machine learning, probabilistic reasoning, natural language processing, knowledge representation and logic. After a brief introduction to the language, programming assignments will l be in Python.
One-term course offered either term

Prerequisite(s): CIS 121, ESE 301 or STAT 430 recommended

ECON 103
Statistics for Economics

The course focuses on elementary probability and inferential statistical techniques. The course begins with a survey of basic descriptive statistics and data sources and then covers elementary probability theory, sampling, estimation, hypothesis testing, correlation, and regression. The course focuses on practical issues involved in the substantive interpretation of economic data using the techniques of statistical inference. For this reason empirical case studies that apply the techniques to real-life data are stressed and discussed throughout the course, and students are required to perform several statistical analyses of their own.

Prerequisites: ECON 001, 002, or ECON 010; MATH 104, 114 or MATH 115

ECON 104
Econometrics

This course is designed to introduce students to econometric techniques and their applications in economic analysis and decision-making. The main objective of the course is to train the student in (i) handling economic data; (ii) quantitative analyses of economic models with probabilistic tools; (iii) econometric techniques, their application as well as their statistical and practical interpretation; (iv) implementing these techniques on a computer. Estimation and inference procedures are formally analyzed for simple econometric models and illustrated by empirical case studies using real-life data. The course covers linear regression models, simultaneous-equations models, discrete choice models and univariate time series models. Estimation and Inference is conducted using least squares and likelihood based techniques. Students are required to perform several econometric analyses of their own.

Prerequisites: ECON 101, ECON 103; MATH 104, MATH 114 or MATH 115 or permission from instructor

ECON 222
Advanced Econometric Techniques and Applications

This course introduces students to advanced study in econometrics, with an emphasis on methods used in microeconomic applications and in evaluating the effects of social interventions. The methods covered include methods for handling limited dependent variables (useful, for example, in forecasting the demand for a new good), maximum likelihood estimators, and flexible semiparametric and non-parametric estimation methods, and randomized and nonexperimental methods of estimating treatment effects. Applications of econometrics to the field of program evaluation will also be studied.

Prerequisites: ECON 101, 104; MATH 104, 114 or MATH 115

ECON 244
Macro-Modeling

This is an advanced undergraduate course in models of economic growth. Students will be introduced to the workhorse theoretical models that are used to understand growth by modern macroeconomic researchers and policy makers. The types of questions that we will address include: Why are some countries richer than others? Why do some countries grow quickly while others stagnate? Why did modern economic growth start in Western Europe? What can governments do to accelerate economic growth? How does economic growth interact with demographic and geographic factors? We will build theoretical models that can be used to answer these questions. There will be a strong focus on emphasizing the microeconomic foundations of models, and using the language of mathematics to express the underlying assumptions and assess their implications for policy. Hence, there are strict mathematical prerequisites. We will also compare the predictions of our models with the data. Thus, a fair amount of econometrics will be required. A class in statistics and econometrics is highly recommended.

Prerequisites: ECON 101 and ECON 102; MATH 104, and MATH 114 or MATH 115. ECON 103 recommended.

ENM 321
Engineering Statistics

This is a first course in applied statistics and probability for students in engineering. The course covers basic concepts of probability, discrete and continuous random variables, probability distributions, data description techniques, random samples, estimations, hypothesis testing, regression, and statistical quality control.

ENM 503
Introduction to Probability and Statistics

Introduction to combinatorics: the multiplication rule, the pigeonhole principle, permutations, combinations, binomial and multinomial coefficients, recurrence relations, methods of solving recurrence relations, permutations and combinations with repetitions, integer linear equation with unit coefficients, distributing balls into urns, inclusion-exclusion, an introduction to probability. Introduction to Probability: sets, sample sets events, axioms of probability, simple results, equally likely outcomes, probability as a continuous set function and probability as a measure of belief, conditional probability, independent events, Bayes' formula, inverting probability trees. Random Variables: discrete and continuous, expected values, functions of random variables, variance. Some Special Discrete Random Variables: Bernoulli, Binomial, Poisson, Geometric, Pascal (Negative Binomial) Hypergeometric and Poisson. Some Special Continuous Random Variables: Uniform, Exponential, Gamma, Erlang, Normal, Beta and Triangular. Joint distribution functions, minimum and maximum of independent random variables, sums of independent random variables, reproduction properties. Properties of Expectation: sums of random variables, covariance, variance of sums and correlations, moment-generating function. Limit theorems: Chebyshev's inequality, law of large numbers and the central-limit theorem. Extra Topics: Generating random numbers and simulation, Monte-Carlo methods, The Poisson Process and Queueing Theory, Stochastic Processes and Regular Markov Chains, Absorbing Markov Chains and Random Walks.

Prerequisite: MATH 240 or equivalent

ESE 301
Engineering Probability

This course introduces students to the mathematical foundations of the theory of probability and its rich applications. The course begins with an exploration of combinatorial probabilities in the classical setting of games of chance, proceeds to the development of an axiomatic, fully mathematical theory of probability, and concludes with the discovery of the remarkable limit laws and the eminence grise of the classical theory, the central limit theorem. The topics covered include: discrete and continuous probability spaces, distributions, mass functions, densities; conditional probability; independence; the Bernoulli schema: the binomial, Poisson, and waiting time distributions; uniform, exponential, normal, and related densities; expectation, variance, moments; conditional expectation; generating functions, characteristic functions; inequalities, tail bounds, and limit laws. But a bald listing of topics does not do justice to the subject: the material is presented in its lush and glorious historical context, the mathematical theory buttressed and made vivid by rich and beautiful applications drawn from the world around us.
The student will see surprises in election-day counting of ballots, a historical wager the sun will rise tomorrow, the folly of gambling, the sad news about lethal genes, the curiously persistent illusion of the hot hand in sports, the unreasonable efficacy of polls and its implications to medical testing, and a host of other beguiling settings.

Prerequisite: MATH 114

ESE 302
Engineering Applications of Statistics

Principles and engineering applications of statistical inference. The basic topics covered are parameter estimation, confidence intervals, and hypothesis testing. Additional topics may include analysis of variance (ANOVA) and/or linear regression. Each method is treated both from theoretical and applied viewpoints, including software analysis of selected data sets.

Prerequisite: ESE 301 or equivalent course in Probability

ESE 303
Stochastic Systems Analysis and Simulation

Stochastic systems analysis and simulation (ESE 303) is a class that explores stochastic systems which we could loosely define as anything random that changes in time. Stochastic systems are at the core of a number of disciplines in engineering, for example communication systems and machine learning. They also find application elsewhere, including social systems, markets, molecular biology and epidemiology. The goal of the class is to learn how to model, analyze and simulate stochastic systems. With respect to analysis we distinguish between what we could call theoretical and experimental analysis. By theoretical analysis we refer to a set of tools which let us discover and understand properties of the system. These analyses can only take us so far and is usually complemented with numerical analysis of experimental outcomes. Although we use the word experiment more often than not we simulate the stochastic system in a computer and analyze the outcomes of these virtual experiments. The class's material is divided in four blocks respectively dealing with Markov chains, continuous time Markov chains, Gaussian processes and stationary processes. Emphasis is placed in the development of toolboxes to analyze these different classes of processes and on describing their applications to complex stochastic systems in different disciplines. Particular examples include: (i) the problem of ranking web pages by a search engine; (ii) the study of reputation and trust in social networks; (iii) modeling and analysis of communication networks; (iv) the use of queues in the modeling of transportation networks; (v) stochastic modeling and simulation of biochemical reactions and gene networks; (vi) arbitrages, pricing of stocks, and pricing of options through Black-Scholes formula; and (vii) linear filtering of stochastic processes to separate signals of interest from background noise. For more information visit the class's web page at http://alliance.seas.upenn.edu/ ~ese303/wiki/.

Prerequisites: ESE 301 or equivalent and one computer language

ESE 304
Optimization of Systems

Model Building and Linear Programming: Graphical Methods and The Simplex Method, the LINDO and LINGO Computer Packages, Degeneracies, Minimization and the BigM and the Two-Phase Methods, and Goal Programming. Sensitivity Analysis: Geometric and Algebraic Approaches, The Computer and Sensitivity Analysis, The Dual of An LP Problem, The Dual Theorem, Shadow Prices, Complementary Slackness, The Dual Simplex Method, and The Revised Simplex Method. Integer Programming: The Branch and Bound Method, Enumeration Methods, and the Cutting Plane Method. Nonlinear Programming: Review of Differential Calculus, Convex and Concave Functions, Solving NLP Problems with One Variable, Unconstraint Nonlinear Optimization with Several Variables, Lagrange Multipliers and Constraint Nonlinear Optimization with Several Variables, The Kuhn-Tucker Conditions and Quadratic Programming.

Corequisite: MATH 240

LGST 242/642
Big Data, Big Responsibilities: The Law and Ethics of Business Analysis

Significant technologies always have unintended consequences, and their effects are never neutral. A world of ubiquitous data, subject to ever more sophisticated collection, aggregation, and analysis, creates massive opportunities for both financial gain and social good. It also creates dangers in areas such as privacy, security, discrimination, exploitation, and inequality, as well as simple hubris about the effectiveness of management by algorithm. Firms that anticipate the risks of these new practices will be best positioned to avoid missteps. This course introduces students to the legal, policy, and ethical dimensions of big data, predictive analytics, and related techniques. It then examines responses-both private and governmental-that may be employed to address these concerns.

*Approved course in the Business Analytics MBA major

MATH 114
Calculus, Part II

Functions of several variables, vector-valued functions, partial derivatives and applications, double and triple integrals, conic sections, polar coordinates, vectors and vector calculus, first order ordinary differential equations. Applications to physical sciences. Use of symbolic manipulation and graphics software in calculus.

Prerequisite: MATH 104

MATH 116
Honors Calculus

Students who are interested in math or science might also want to consider a more challenging Honors version of Calculus II and III, Math 116 and Math 260 (the analogues of Math 114 and Math 240, respectively). These courses will cover essentially the same material as 114 and 240, but more in depth and involve discussion of the underlying theory as well as computations.

MATH 240
Calculus, Part III

Linear algebra: vectors, matrices, systems of linear equations, vector spaces, subspaces, spans, bases, and dimension, eigenvalues, and eigenvectors, matrix exponentials. Ordinary differential equations: higher-order homogeneous and inhomogeneous ODEs and linear systems of ODEs, phase plane analysis, non-linear systems.

Prerequisite: Calculus II (MATH 114)

MATH 241
Calculus, Part IV

Partial differential equations and their solutions, including solutions of the wave, heat and Laplace equations, and Sturm-Liouville problems. Introduction to Fourier series and Fourier transforms. Computation of solutions, modeling using PDE's, geometric intuition, and qualitative understanding of the evolution of systems according to the type of partial differential operator.

Prerequisite: MATH 240

MATH 312/412
Linear Algebra

Linear transformations, Gauss Jordan elimination, eigenvalues and eigenvectors, theory and applications. Mathematics majors are advised that MATH 312 cannot be taken to satisfy the major requirements.

Prerequisite(s): MATH 240

Notes: Students who have already received credit for either Math 370, 371, 502 or 503 cannot receive further credit for Math 312 or Math 313/513. Students can receive credit for at most one of Math 312 and Math 313/513

MATH 313/513
Computational Linear Algebra

Many important problems in a wide range of disciplines within computer science and throughout science are solved using techniques from linear algebra. This course will introduce students to some of the most widely used algorithms and illustrate how they are actually used. Some specific topics: the solution of systems of linear equations by Gaussian elimination, dimension of a linear space, inner product, cross product, change of basis, affine and rigid motions, eigenvalues and eigenvectors, diagonalization of both symmetric and non-symmetric matrices, quadratic polynomials, and least squares optimization. Applications will include the use of matrix computations to computer graphics, use of the discrete Fourier transform and related techniques in digital signal processing, the analysis of systems of linear differential equations, and singular value decompositions with application to a principal component analysis. The ideas and tools provided by this course will be useful to students who intend to tackle higher level courses in digital signal processing, computer vision, robotics, and computer graphics.

Prerequisites: Math 240, and some programming experience.

Notes: Students who have already received credit for either Math 370, 371, 502 or 503 cannot receive further credit for Math 312 or Math 313. Students can receive credit for at most one of Math 312 and Math 313.

MATH 530
Mathematics of Finance

This course presents the basic mathematical tools to model financial markets and to make calculations about financial products, especially financial derivatives. Mathematical topics covered: stochastic processes, partial differential equations and their relationship. No background in finance is assumed.

Prerequisites: MATH 240, STAT 430

MKTG 101
Introduction to Marketing

The objective of this course is to introduce students to the concepts, analyses, and activities that comprise marketing management, and to provide practice in assessing and solving marketing problems. The course is also a foundation for advanced electives in Marketing as well as other business/social disciplines. Topics include marketing strategy, customer behavior, segmentation, market research, product management, pricing, promotion, sales force management and competitive analysis.

MKTG 212/712
Data and Analysis for Marketing Decisions

Firms have access to detailed data of customers and past marketing actions. Such data may include in-store and online customer transactions, customer surveys as well as prices and advertising. Using real-world applications from various industries, the goal of the course is to familiarize students with several types of managerial problems as well as data sources and techniques, commonly employed in making effective marketing decisions. The course would involve formulating critical managerial problems, developing relevant hypotheses, analyzing data and, most importantly, drawing inferences and telling convincing narratives, with a view of yielding actionable results.

Prerequisite(s): MKTG 101, STAT 101. Students are highly encouraged to take statistics in the semester immediately preceding this course. (Former course title Marketing Research.)

*Approved course in the Business Analytics MBA major

MKTG 271/771
Models for Marketing Strategy

In today's business environment, marketing executives are involved in complex decision-making and they become responsible for return on their marketing investments. The first objective of this course is to help participants become better executives. By exposing students to various analytical and computer-based tools, developed for solving marketing problems, it will help to prepare them for careers in industries such as consumer packaged goods, hi-tech, financial services, media and entertainment, pharmaceutical, consulting, and venture capital. The course's main focus is on various existing models, such as models that predict the consumer's dynamic adoption of an innovative product. However, at some point in their career, students may find themselves facing business problems for which a model can assist in making decisions, but no existing model is available. Hence, the second objective of the course is to provide participants with critical skills necessary to evaluate new models to which they may be exposed by attending presentations or reading the literature. The models to be discussed in the class have been implemented and proven useful in a wide range of industries (e.g., business-to-consumers and business-to-business). The course is not only about models, however. It also covers modeling needs. Some industries such as the media and entertainment or the pharmaceutical industries present unique problems and modeling needs. The third objective of the course is to expose participants to the nature and essence of such idiosyncratic problems as well as modeling needs in such industries. Overall, the course will make participants understand better critical marketing problems by analyzing them rigorously and will enhance their skills in either designing or evaluating models-based strategies.

Taught by: Eliashberg

Prerequisites: MKTG 101, STAT 101 and some tolerance for expressing critical ideas in simple math

*Approved course in the Business Analytics MBA major

MKTG 476/776 / STAT 476/ 776
Applied Probability Models for Marketing

This course will expose students to the theoretical and empirical "building blocks" that will allow them to construct, estimate, and interpret powerful models of customer behavior. Over the years, researchers and practitioners have used these models for a wide variety of applications, such as new product sales, forecasting, analyses of media usage, and targeted marketing programs. Other disciplines have seen equally broad utilization of these techniques. The course will be entirely lecture-based with a strong emphasis on real-time problem solving. Most sessions will feature sophisticated numerical investigations using Microsoft Excel. Much of the material is highly technical.

Taught by: Fader

Prerequisites: Prerequisite(s): A high comfort level with basic integral calculus and recent exposure to a formal course in probability and statistics such as STAT 430 is strongly recommended.

*Approved course in the Business Analytics MBA major

MKTG 309/809
Experiments for Business Decision Making

In the past decade, massive shifts in how companies interact with their customers have suddenly made field experiments an economically feasible way to learn about a variety of business questions such as what types of promotions are most effective, what products should be stocked at a store, how e-mail promotions should be designed, how sales staff should be compensated, etc. Many marketers engaged in online retailing, direct-marketing, online advertising, media management, etc. are rapidly embracing a "test and learn" philosophy and a number of platforms such as Google Website Optimizer, have been developed to facilitate rigorous field experiments in the online environment. Just as with the quality revolution in manufacturing during the 1980s and 1990s, the rapid rise of the "test and learn" philosophy in marketing has created a huge demand for those who can design, field, and analyze marketing experiments. Through this course, you will learn and practice a wide range of critical skills, from the statistical methods used to design and analyze experiments to the management and strategy required to execute an experiment and act on the results. Although the cases and examples will focus on marketing problems, the material covered can be applied in a number of other domains particularly operations management and product design.

Prerequisites: MKTG 101 or faculty permission is required; STAT 101, STAT 431, or equivalent is recommended. (Former MKTG 269)

Course not offered every year

*Approved course in the Business Analytics MBA major

MKTG 940/941
Measurement and Data Analysis in Marketing, Parts A and B

In this course we consider models for binary, count, and continuous data including contingency table models, logistic and profit regression, ANOVA, ANCOVA, conjoint analysis, and OLS. In addition we cover multidimensional techniques such as MDS, cluster analysis, principal components analysis, factor analysis, and discriminant analysis. We utilize the statistics package SPlus 2000, and also BUGS for implementing many of the techniques described in a Bayesian manner.

OIDD 101
An Introduction to Operations, Information, and Decisions

OIDD 101 explores a variety of common quantitative modeling problems that arise frequently in business settings, and discusses how they can be formally modeled and solved with a combination of business insight and computer-based tools. The key topics covered include capacity management, service operations, inventory control, structured decision making, constrained optimization and simulation. This course teaches how to model complex business situations and how to master tools to improve business performance. The goal is to provide a set of foundational skills useful for future coursework at Wharton as well as providing an overview of problems and techniques that characterize disciplines that comprise Operations and Information Management.

OIDD 105
Developing Tools for Data Access and Analysis (VBA and SQL Programming)

This course provides an introduction to the construction of data analysis tools that are commonly used for business applications, especially in consulting and finance. The course builds on the spreadsheet and analytical skills developed in OPIM101, providing a much more extensive treatment of spreadsheet application development and database management. The first portion of the course will focus on programming in VBA, the embedded programming language in the Microsoft Office suite of applications. This will be supplemented with discussion of industry best practice in software development, such as specification development, interface design, documentation, and testing. The second portion of the class will emphasize data access and analysis utilizing SQL, the industry standard language for interacting with database software.

OIDD 290
Decision Processes

This course is an intensive introduction to various scientific perspectives on the processes through which people make decisions. Perspectives covered include cognitive psychology of human problem-solving, judgment and choice, theories of rational judgment and decision, and the mathematical theory of games. Much of the material is technically rigorous. Prior or current enrollment in STAT 101 or the equivalent, although not required, is strongly recommended.

Prerequisite: STAT 101 or equivalent strongly recommended

OIDD 311
Business Computer Languages

This course is taught with the more descriptive title of "Scripting for Business Analytics." "Business Analytics" refers to modeling and analysis undertaken for purposes of management and supporting decision-making. The varieties of techniques and methods are numerous and growing, including simple equational models, constrained optimization models, probabilistic models, visualization, data analysis, and much more. Elementary modeling of this sort can be undertaken in Excel and other spreadsheet programs, but "industrial strength" applications typically use more sophisticated tools, based on scripting languages. Scripting languages are programming languages that are designed to be learned easily and to be used for special purposes, rather than for large-scale application programming. This course focuses on the special purposes associated with business analytics and teaches MATLAB and Python in this context. MATLAB and Python are widely used in practice (both in management and in engineering), as are the business analytic methods covered in the course. Prior programming experience is useful, but not required or presumed for this course.

OIDD 314/662
Enabling Technologies

Conducting business in a networked economy invariably involves interplay with technology. The purpose of this course is to improve understanding of technology (what it can or cannot enable) and the business drivers of technology-related decisions in firms. We will be discussing some of the new and most disruptive technologies right now to stimulate thought on new applications for commerce and new ventures, as well as their implications to the tech industry as a whole. Topics include social media, online advertising, big data, and cloud computing. The course will take a layered approach (from network infrastructure) to data infrastructure to applications infrastructure, or direct enablers of commerce) to first, understanding and then, thinking about technology enablers. Network infrastructure layers include fundamentals of wired and wireless infrastructure technologies such as protocols for networking, broadband technologies - for last (DSL, Cable etc.) and other miles (advances in optical networking) and digital cellular communications. Data infrastructure layers include usage tracking technologies, search technologies and data mining. Direct application layers include personalization technologies (CRM), design technologies for content and exchanges, software renting enablers, application service provision, agents and security mechanisms. Finally some emerging technology enablers (such as Bluetooth, biometrics and virtual reality) are identified and discussed.

OIDD 321
Introduction to Management Science

Understanding how to use data and business analytics can be the key differential for a company's success or failure. This course is designed to introduce fundamental quantitative decision making tools for a broad range of managerial decision problems. Topics covered include linear, nonlinear and discrete optimization, dynamic programming, and simulation. Students will apply these quantitative models in applications of portfolio management, electricity auctions, revenue management for airlines, manufacturing, advertising budget allocation, and healthcare scheduling operations. Emphasis in this course is placed on mathematical modeling of real world problems and implementation of decision-making tools.

OIDD 325/PHIL 203
Computer Simulation Models

This course focuses on agent-based computational models in the social sciences, especially in economic, in commercial and in strategic (game-theoretic) contexts. This relatively recent and now rapidly-developing form of computer simulation seeks to explain and predict complex social phenomena "from the ground up", through interactions of comparatively simple agents. The course reviews experimental and theoretical results, and exposes the students to modern development environments for this form of simulation. Students have the opportunity to design and implement agent-based simulations. Programming, however, is not required. This course aims to integrate various topics in agent-based simulation, while developing an appreciation of the problems that are particularly characteristic of this form of simulation so that students will understand its promise and potential.

OIDD 353/653
Mathematical Modeling and its Application in Finance

Quantitative methods have become fundamental tools in the analysis and planning of financial operations. There are many reasons for this development: the emergence of a whole range of new complex financial instruments, innovations in securitization, the increased globalization of the financial markets, the proliferation of information technology and the rise of high-frequency traders, etc. In this course, models for hedging, asset allocation, and multi-period portfolio planning are developed, implemented, and tested. In addition, pricing models for options, bonds, mortgage-backed securities, and other derivatives are studied. The models typically require the tools of statistics, optimization, and/ or simulation, and they are implemented in spreadsheets or a high-level modeling environment, MATLAB. This course is quantitative and will require extensive computer use. The course is intended for students who have strong interest in finance. The objective is to provide students the necessary practical tools they will require should they choose to join the financial services industry, particularly in roles such as: derivatives, quantitative trading, portfolio management, structuring, financial engineering, risk management, etc. Prospective students should be comfortable with quantitative methods such as basic statistics and the methodologies (mathematical programming and simulation) taught in OIDD 612 Business Analytics and OIDD 321 Management Science (or equivalent). Students should seek permission from the instructor if the background requirements are not met.

Prerequisite(s): OPIM321 (or a similar optimization class).

*Approved course in the Business Analytics MBA major

OIDD 410/672
Decision Support Systems

The past few years have seen an explosion in the amount of data collected by businesses and have witnessed enabling technologies such as database systems, client-server computing and artificial intelligence reach industrial strength. These trends have spawned a new breed of systems that can support the extraction of useful information from large quantities of data. Understanding the power and limitations of these emerging technologies can provide managers and information systems professionals' new approaches to support the task of solving hard business problems. This course will provide an overview of these techniques (such as genetic algorithms, neural networks, and decision trees) and discuss applications such as fraud detection, customer segmentation, trading, marketing strategies and customer support via cases and real datasets.

*Approved course in the Business Analytics MBA major

OIDD 642
Analytics for Services

This course covers a range of analytical methods that are useful tools for capacity management in services, and it will provide you with insights into the economics of a range of services businesses including (i) High-level planning models that account for multiple dimensions of service capacity, (ii) Low-level models of system congestion that capture the relationship between capacity choices, quality of service and, in some cases, system revenue, (iii) Statistical estimation and forecasting models to characterize key measures of future supply and demand.

Prerequisite(s): Students who have already taken OIDD 611, OIDD 612, and STAT 613 should be well-equipped for the class. Other students should have a solid understanding of elementary probability, statistics and linear programming. For questions regarding the specifics of your background, please contact the instructor.

*Approved course in the Business Analytics MBA major

OIDD 643
Analytics for Revenue Management

This course introduces you to the essential concepts and techniques required to understand and implement revenue management (RM). The need for repeated, rapid and cycles of estimation and optimization has driven the development of a set of analytical tools that are particularly well suited for RM. This course focuses on those tools.

Prerequisites: Students who have already taken OPIM 612 and STAT 613 should be well equipped for this class. Other students should have a solid understanding of elementary probability, statistics and constrained optimization. For questions regarding the specifics of your background, please contact the instructor.

*Approved course in the Business Analytics MBA major

OIDD 658
Service Operations Management

The service sector represents the largest segment of most industrial economies. In the U.S., for example, it accounts for approximately 70% of GDP and 70% of employment. In addition to this "pure" service sector, the operations and competitive positions of many manufacturing firms are becoming increasingly service-oriented. While operational excellence is critical for success in most industries today, in a wide range of service industries this is particularly true. For example, recent, significant deregulation in banking, health care, and communications has led to intensified competition and pressure on operations. At the same time, the rapid evolution of information technology has enabled firms to operate in a fashion - and offer a level of service - that has not been previously possible. Elements common to most services make the management of their operations complex, however. In particular, services are intangible, not storable or transportable, and often highly variable. Frequently their delivery involves distributed operations with a significant amount of customer contact. All of these factors make service operations end up looking quite a bit different than manufacturing operations, and the task of achieving excellence in them requires specialized analysis frameworks and tools. This course covers a mix of qualitative and quantitative models that provide the necessary tools. The class will focus on simple models that should help you to better understand both the difficulty of managing and the underlying economics of the service operations being considered. You will have the opportunity to apply these course tools in a group service assessment field project.

Prerequisite(s): Courses in operations management, linear programming, probability and statistics

*Approved course in the Business Analytics MBA major

OIDD 664
Data Based and Information Management Systems

Data and information are critical to the modern organization. Whether used in knowledge management, business intelligence, enterprise resource planning (ERP), product design, marketing, personalization and other aspects of managing customer relationships (CRM), the underlying principles of data management are the same. This course aims to provide a practical introduction to the fundamental principles. Examples and exercises will cover the relational database tools at the core of ERP, CRM, and on-line exchanges and portals. However, the course will also use the same basic foundations to consider emerging technologies and standards such as XML, ebXML, UDDI, etc.

OID 898
Advanced Topics

This course will show how to use Big Data from the internet to better understand and predict behavior. We will discuss how to make sense of Google searches, social media posts and other new sources of data. We will particularly emphasize the importance of finding honest sources of data, since people often lie. Students will learn both to make sense of existing studies and to conduct their own studies. The course is taught by Dr. Seth Stevens-Davidowitz.

*Approved course in the Business Analytics MBA major

OIDD 910/ESE 504
Concepts of Math Programming

Introduction to mathematical programming for PhD students who would like to be intelligent and sophisticated consumers of mathematical programming theory but do not plan to specialize in this area. Integer and nonlinear programming are covered, including the fundamentals of each area together with a sense of the state-of-the-art and expected directions of future progress

OIDD 930
Stochastic Models

This course introduces mathematical models describing and analyzing the behavior of processes that exhibit random components. The theory of stochastic processes will be developed based on elementary probability theory and calculus. Topics include random walks, Poisson processes, Markov chains in discrete and continuous time, renewal theory, and martingales. Applications from the areas of inventory, production, finance, queueing and communication systems will be presented throughout the course. This course is really only appropriate for students with a very strong math background (i.e. a semester of linear algebra, a semester of real analysis, a semester of advanced probability, and exposure to differential equations) and an interest in going into academia.

Prerequisite: STAT 510 or equivalent

OIDD 931/STAT 901
Stochastic Processes II

Extension of the material presented in OIDD930 to include renewal theory, martingales, and Brownian motion.

Prerequisite: OPIM 930

OIDD 934
Dynamic Programming and Stochastic Models

The course goal is to provide a brief but fairly rigorous introduction to the formulation and solution of dynamic programs. Its focus is primarily methodological. We will cover discrete state space problems, over finite or infinite time horizon, with and without discounting. Structured policies and their theoretical foundation will be of particular interest. Computational methods and approximation methods will be addressed. Applications are presented throughout the course, such as inventory policies, production control, financial decisions, and scheduling.

Prerequisite: OPIM 930

PSYC 739
Special Topics in Perception

Probability theory has become an increasingly popular and successful framework for modeling human perceptual and cognitive behavior. This course will provide a careful introduction to probability theory and the various ways it has been applied in psychology and neuroscience. Goal is to make students understand the most important state-of-the-art probabilistic models in perception and cognition, what they reveal about the brain's underlying computations and strategies in dealing with uncertainty, and how such computations can potentially be performed by populations of neurons.

Taught by: Stocker

STAT 101/102
Introductory Business Statistics

101: Data summaries and descriptive statistics; introduction to a statistical computer package; Probability: distributions, expectation, variance, covariance, portfolios, central limit theorem; statistical inference of univariate data; Statistical inference for bivariate data: inference for intrinsically linear simple regression models. This course will have a business focus, but is not inappropriate for students in the college.

102: Continuation of STAT 101. A thorough treatment of multiple regression, model selection, analysis of variance, linear logistic regression; introduction to time series. Business applications.

Prerequisites: MATH 104 or equivalent; successful completion of STAT 101 is prerequisite to STAT 102

STAT 111/112
Introductory Statistics

111: Introduction to concepts in probability. Basic statistical inference procedures of estimation, confidence intervals and hypothesis testing directed towards applications in science and medicine. The use of the JMP statistical package. Prerequisite: High school algebra

112: Further development of the material in STAT 111, in particular the analysis of variance, multiple regression, non-parametric procedures and the analysis of categorical data. Data analysis via statistical packages. Successful completion of STAT 111 is a prerequisite to STAT 112

STAT 405/705
Statistical Computing with R

The goal of this course is to introduce students to the R programming language and related ecosystem. This course will provide a skill-set that is in demand in both the research and business environments. In addition, R is a platform that is used and required in other advanced classes taught at Wharton, so that this class will prepare students for these higher level classes and electives.

Taught by: Stine, Waterman, Zhang

Prerequisite: STAT 102 or STAT 112 or STAT 430

0.5 CU course

*Approved course in the Business Analytics MBA major

STAT 422/722
Predictive Analysis for Business

This course follows from the introductory regression classes, STAT 102, STAT 112, and STAT 431 for undergraduates and STAT 613 for MBAs. It extends the ideas from regression modeling, focusing on the core business task of predictive analytics as applied to realistic business related data sets. In particular it introduces automated model selection tools, such as stepwise regression and various current model selection criteria such as AIC and BIC. It delves into classification methodologies such as logistic regression. It also introduces classification and regression trees (CART) and the popular predictive methodology known as the random forest. By the end of the course the student will be familiar with and have applied all these tools and will be ready to use them in a work setting. The methodologies can all be implemented in either the JMP or R software packages.

Prerequisite: STAT 102 or STAT 112 or STAT 430

0.5 CU course

STAT 430/510
Probability

Discrete and continuous sample spaces and probability; random variables, distributions, independence; expectation and generating functions; Markov chains and recurrence theory.

Prerequisite: MATH 114 or MATH 115 or equivalent

STAT 431 /511
Statistical Inference

Graphical displays; one- and two-sample confidence intervals; one- and two-sample hypothesis tests; one- and two-way ANOVA; simple and multiple linear least-squares regression; nonlinear regression; variable selection; logistic regression; categorical data analysis; goodness-of-fit tests. A methodology course. This course does not have business applications but has significant overlap with STAT 101 and 102.

Prerequisite: STAT 430

STAT 433/533
Stochastic Processes

An introduction to Stochastic Processes. The primary focus is on Markov Chains, Martingales and Gaussian Processes. We will discuss many interesting applications from physics to economics. Topics may include: simulations of path functions, game theory and linear programming, stochastic optimization, Brownian Motion and Black-Scholes.

Taught by: Steele, Mossel
Prerequisite: STAT 430, or permission of instructor

STAT 434
Financial and Economic Time Series

This course will introduce students to the time series methods and practices, which are most relevant to the analysis of financial and economic data. After an introduction to the statistical programming language R the course develops an autoregressive models, moving average models, and their generalizations. The course then develops models that are closely focused on particular features of financial series such as the challenges of time dependent volatility.

Prerequisites: STAT 101-102 or 431, familiarity with linear algebra

STAT 435/711
Forecasting Methods for Management

This course provides an introduction to the wide range of techniques available for statistical forecasting. Qualitative techniques, smoothing and decomposition of time series, regression, adaptive methods, autoregressive-moving average modeling, and ARCH and GARCH formulations will be surveyed. The emphasis will be on applications, rather than technical foundations and derivations. The techniques will be studied critically, with examination of their usefulness and limitations.

Prerequisite: STAT 102 or 112 or 431

*Approved course in the Business Analytics MBA major

STAT 453/853 / BEPP 453/853
Actuarial Statistics

This course covers models for insurer's losses, and applications of Markov chains. Poisson processes, including extensions such as non-homogeneous, compound, and mixed Poisson processes are studied in detail. The compound model is then used to establish the distribution of losses. An extensive section on Markov chains provides the theory to forecast future states of the process, as well as numerous applications of Markov chains to insurance, finance, and genetics. The course is abundantly illustrated by examples from the insurance and finance literature. While most of the students taking the course are future actuaries, other students interested in applications of statistics may discover in class many fascinating applications of stochastic processes and Markov chains.

Taught by: Jean Lemaire

Prerequisite: STAT 430

STAT 470/503/770
Data Analytics and Statistical Computing

This course will introduce a high-level programming language, called R, that is widely used for statistical data analysis. Using R, we will study and practice the following methodologies: data cleaning, feature extraction; web scrubbing, text analysis; data visualization; filing statistical models; simulation of probability distributions and statistical models; statistical inference methods that use simulations (bootstrap, permutation tests).

Taught by: Buja

Prerequisites: STAT 101 and 102 OR STAT 111 and 112 OR STAT 431

*Approved course in the Business Analytics MBA major

STAT 471/571/701
Modern Data Mining

Statistics or Data Science has been evolving rapidly to keep up with the modern world. While classical multiple regression and logistic regression technique continue to be the major tools we go beyond to include methods built on top of linear models such as LASSO and Ridge regression. Contemporary methods such as KNN (K nearest neighbor), Random Forest, Support Vector Machines, Principal Component Analyses (PCA), the bootstrap and others are also covered. Text mining especially through PCA is another topic of the course. While learning all the techniques, we keep in mind that our goal is to tackle real problems. Not only do we go through a large collection of interesting, challenging real-life data sets but we also learn how to use the free, powerful software "R" in connection with each of the methods exposed in the class.

Taught by: Linda Zhao

Prerequisite: STAT 613 or equivalent

*Approved course in the Business Analytics MBA major

STAT 472/712
Decision Making Under Uncertainty

Fundamentals of modern decision analysis with emphasis on managerial decision making under uncertainty and risk. The basic topics of decision analysis are examined. These include payoffs and losses, utility and subjective probability, the value of information, Bayesian analysis, inference and decision making. Examples are presented to illustrate the ideas and methods. Some of these involve: choices among investment alternatives; marketing a new product; health care decisions; and costs, benefits, and sample size in surveys.

Taught by: Stine

Prerequisite: STAT 102 or 112 or 431

STAT 473/953
Bioinformatics

An introduction to the use of statistical methods in the increasingly important scientific areas of genomics and bioinformatics. The topics to be covered will be decided in detail after the initial class meeting, but will be taken from the following: - background probability theory of one and many random variables and of events; background statistical inference theory, classical and Bayesian; Poisson processes and Markov chain; the analysis of one and many DNA sequences, in particular shotgun sequencing, pattern analysis and motifs; substitution matrices, general random walk theory, advanced statistical inference, the theory of BLAST, hidden Markov models, microarray analysis, evolutionary models.

Taught by: Ewens

Prerequisites: Good background in probability and statistics at the approximate level of STAT 430 and STAT 431.

STAT 474/CRIM 474/STAT 974
Modern Regression

Function estimation and data exploration using extensions of regression analysis: smoothers, semiparametric and nonparametric regression, and supervised machine learning. Conceptual foundations are addressed as well as hands-on use for data analysis.

Taught by: Berk

Prerequisite: STAT 102 or 112 or equivalent

STAT 475/BSTA 775/STAT 920
Sample Survey Design

This course will cover the design and analysis of sample surveys. Topics include simple sampling, stratified sampling, cluster sampling, graphics, regression analysis using complex surveys and methods for handling nonresponse bias.

Prerequisite: STAT 102 or 112 or 431

STAT 520/521
Applied Econometrics I/II

520: This is a course in econometrics for graduate students. The goal is to prepare students for empirical research by studying econometric methodology and its theoretical foundations. Students taking the course should be familiar with elementary statistical methodology and basic linear algebra, and should have some programming experience. Topics include conditional expectation and linear projection, asymptotic statistical theory, ordinary least squares estimation, the bootstrap and jackknife, instrumental variables and two-stage least squares, specification tests, systems of equations, generalized least squares, and introduction to use of linear panel data models.

521: Topics include system estimation with instrumental variables, fixed effects and random effects estimation, M-estimation, nonlinear regression, quantile regression, maximum likelihood estimation, generalized method of moments estimation, minimum distance estimation, and binary and multinomial response models. Both theory and applications will be stressed. Successful completion of STAT 520 is a prerequisite to STAT 521.

Taught by: Shaman

Prerequisites: MATH 114 and MATH 312 or equivalents, and an undergraduate introduction to probability and statistics.

STAT 542
Bayesian Methods and Computation

Sophisticated tools for probability modeling and data analysis from the Bayesian perspective. Hierarchical models, mixture models and Monte Carlo simulation techniques.

Taught by: Shane Jensen

Prerequisites: STAT 430 or 510 or equivalent or permission of instructor

STAT 422/722
Predictive Analytics for Business

This course follows from the introductory regression classes, STAT 102, STAT 112, and STAT 431 for undergraduates and STAT 613 for MBAs. It extends the ideas from regression modeling, focusing on the core business task of predictive analytics as applied to realistic business related data sets. In particular it introduces automated model selection tools, such as stepwise regression and various current model selection criteria such as AIC and BIC. It delves into classification methodologies such as logistic regression. It also introduces classification and regression trees (CART) and the popular predictive methodology known as the random forest. By the end of the course the student will be familiar with and have applied all these tools and will be ready to use them in a work setting. The methodologies can all be implemented in either the JMP or R software packages.

Prerequisite(s): STAT 613 or STAT 621 or having waived the statistics core completely.

*Approved course in the Business Analytics MBA major

STAT 424/724
Text Analytics

This course introduces methods for the analysis of unstructured data, focusing on statistical models for text. Techniques include those for sentiment analysis, topic models, and predictive analytics. Course includes topics from natural language processing (NLP), such as identifying parts of speech, parsing sentences (e.g., subject and predicate), and named entity recognition (people and places). Unsupervised techniques suited to feature creation provide variables suited to traditional statistical models (regression) and more recent approaches (regression trees). Examples that span the course illustrate the success of text analytics. Hierarchical generating models often associated with nonparametric Bayesian analysis supply theoretical foundations.

Prerequisites: Students should be familiar with regression models at the level of STAT 613 and the R statistics language at the level of STAT 705. Familiarity with the R-Studio development environment is presumed, as well as common R packages such as stringr, dplyr and ggplot. Those with more knowledge of Statistics, such as from STAT 722, or computing skills will benefit. The predominant software used in the course is R, with bits of JMP when helpful for interactive illustration. Familiarity with basic probability models is helpful but not presumed.

*Approved course in the Business Analytics MBA major

STAT 901/OIDD 931
Stochastic Processes II

Martingales, optimal stopping, Wald's lemma, age-dependent branching processes, stochastic integration, Ito's lemma.

Prerequisite: OPIM 930 or equivalent

STAT 955
Stochastic Calculus and Financial Applications

Selected topics in the theory of probability and stochastic processes. Requires a strong background in measure theory and real analysis.

Taught by: Steele

Prerequisite: STAT 930 or equivalent

STAT 961
Statistical Methodology

This is a course that prepares 1st year PhD students in statistics for a research career. This is not an applied statistics course. Topics covered include: linear models and their high-dimensional geometry, statistical inference illustrated with linear models, diagnostics for linear models, bootstrap and permutation inference, principal component analysis, smoothing and cross-validation.

Taught by: Andreas Buja

Prerequisites: STAT 431 or 520 or equivalent; a solid course in linear algebra and a programming language

STAT 970
Mathematical Statistics

Decision theory and statistical optimality criteria, sufficiency, point estimation and hypothesis testing methods and theory.
Prerequisites: STAT 431 or 520 or equivalent; comfort with mathematical proofs (e.g., MATH 360).

Taught by: Dylan Small

Prerequisites: STAT 431 or 520 or equivalent; comfort with mathematical proofs (e.g., MATH 360)