DATA ANALYTICS - DAT

DAT 500 Interactive Graphical Case Studies in Big Data 1 Credit
Students will be introduced to Data Analytics via the study of a variety of
case studies of published studies, or successful commercial applications
of methods. Students will also learn to replicate the graphical presentations
used in these studies, and develop alternative visual representations
of the data used in the studies. The R statistical language will be used, as
students learn how to produce publication grade graphics that can be used
throughout other courses and in their career.
Offered: every summer.

DAT 501 Statistics and Econometrics 3 Credits
Econometrics is the science in which the tools of economic theory,
mathematics and statistical inference are applied to the analysis of economic
phenomena. Econometric modeling is an important research tool in
Economics, Finance, and many other academic disciplines. The goal of this
course is to provide you with a basic understanding of Econometric theory
and practice. We will focus on model specification, estimation, and testing,
using a "hands on" approach. Both EXCEL and EViews software will be used
throughout this course.

DAT 511 Data Stewardship: Preparation, Exploration and Handling of Big
Data 3 Credits
Data stewardship refers to the process of managing collections of data in
an ethical and effective manner, so that business objectives can be achieved
efficiently while respecting the rights of individuals. This course will thus
cover the substantial ethical issues related to Big Data, but will also address
many technical issues related to working with large data sets. Establishing
and maintaining quality data poses surprisingly large challenges and can be
very time consuming, so that knowledge of effective data cleaning is a key
capability for Data Analytics. Students will learn how to download, clean,
and prepare data for future analysis, and document the process, as well
as understanding how seeming harmless actions can pose threats to the
information security of others.
Offered: every fall.

DAT 512 Statistical Approaches to Big Data 3 Credits
This course is a Core course in the Data Analytics program. It starts with a
brief review of univariate statistics and then covers selected topics usually
taught in courses in multivariate statistical analysis and regression analysis.
It is assumed that every student in this course has completed at least one
college-level statistics course. The theoretical knowledge and analytical
skills gained in this course are an essential component of the Data Analytics
program.
Prerequisite: DAT 501 or equivalent.
Offered: every spring.

DAT 513 Database Management 3 Credits
This course presents an introduction to the design and use of database
systems. Traditional databases will be the primary focus, centering on the
relational model (SQL and related tools). There will be some discussion of
large-scale information retrieval in the form of the NoSQL movement and
data mining. Ethical, social and security issues will also be covered in an
introductory fashion.
Prerequisites: CSC 501 and CSC 502, or equivalent.
Offered: every fall.

DAT 514 Data Mining and Machine Learning 3 Credits
This course is a Core course in the Data Analytics program. It starts with a
brief introduction to Data Mining and Statistical Learning, includes a brief
summary of relevant methods covered in a much greater detail in other
courses in this program, such as Data Stewardship and Statistical approaches
to Big Data, and then covers a number of methods essential in the modern
Data Mining and Statistical Learning.
Prerequisites: MAT 500, DAT 511, DAT 512.
Offered: every spring.

DAT 515 Visualization and Presentation of Advanced Analytics 3 Credits
Students will develop the ability to present complex results from Data
Analytics to a range of audiences. The course will cover both real time
interactive displays and tools, such as graphic user interface and dashboard
design, as well as written, oral and graphical communication of analytic
results. Students will complete a range of projects in each of these areas.
Prerequisites: DAT 511, DAT 521, and ability to program in Python.
Offered: every spring.

DAT 521 Applied Integrative Projects in Data Analytics I 2 Credits
In this course, students would learn SAS. Since the focus is on hands-on, all
lectures would be conducted in a computer lab. Students learn how to input
various types of data into SAS, such as text, csv, binary and sas7bdat. How
to clean data is an important skill students are expected to master. Students
learn how to deal with missing variables and run basic sample statistics such
as mean, standard deviation, minimum and maximum. Many visualization
techniques would be taught. In addition, students learn how to run some
basic statistical functions, such as linear regression. Since this course is
a preparation for the next course (DAT 522) titled "Applied Integrative
Projects in Data Analytics II", students could start to think about their next
big projects.
Offered: every fall.

DAT 522 Applied Integrative Projects in Data Analytics II 3 Credits
This course focuses on hands-on and term project. It serves as a link between
many core courses, such as Data Cleaning, Machines Learning and domain
knowledge, such as Economics, Accounting, Finance, and Marketing.
Students would apply what they have learnt, such as machine Learning, to
various real world situations. For students with accounting background,
they learn how to process 10-K (annual reports downloaded from SEC's
web site). For students with a background of Economics, they learn how to
generate SAS and R data sets from the data downloaded from the Federal
Reserve Bank's Data Library and US Census and apply them to predict the
market moments. For students with a finance background, they learn how
to process CRSP and Compustat to evaluate various trading strategies,
such as momentum strategy, industry momentum strategy, 52-high trading
strategy. In addition, they learn how to generate various SAS and R data sets
from Prof. French's Data Library. For students with marketing knowledge,
they learn how to parse social media data to fine tune their marketing
strategies. For students from other areas, they learn how to estimate the
gender and age groups by analyzing million cell phone's usages such as
brand, event, timestamp of the events, app downloaded. This course uses
two languages are SAS and R.
Prerequisites: DAT 500, DAT 514, DAT 521.
Offered: every spring.