DATA ANALYTICS (MS OR ADVANCED CERTIFICATE)

Students may choose to pursue this program full time (cohort based) or part time. A graduate certificate program is also available.

If you are interested in more information about this program, please contact Michael Seaman at 716-888-2545 or seamanm@canisius.edu.

PROGRAM DIRECTOR
H. David Sheets, Professor of Physics (716-888-2587 or sheets@canisius.edu)

PROGRAM FACULTY
Debra T. Burhans, PhD, Associate Professor of Computer Science
Milburn E. Crotzer, MBA, PhD, Adjunct Professor of Mathematics and Statistics
Byung-Jay Kahng, PhD, Professor and Chair of Mathematics and Statistics
Leonid A. Khinkis, PhD, Professor of Mathematics and Statistics
Jonathan E. Lopez, PhD, Assistant Professor of Mathematics and Statistics
Jeffrey J. McConnell, PhD, Professor and Chair of Computer Science
Mark S. Meyer, PhD, Associate Professor of Computer Science
Paul Sauer, MBA, PhD, Professor of Marketing and Information Systems
Richard Wall, CFA, PhD, Professor of Economics and Finance
Michael H. Wood, PhD, Associate Professor and Chair of Physics

Overview
The Masters Program in Data Analytics at Canisius is offered in a full-time cohort system, on campus, and may be completed in one calendar year, or as a part-time program taking two years, depending on preparation. A key feature is the incorporation of Applied Integrative Projects, ideally internships, beginning early in the program and paralleling advanced coursework. A 4+1 program for students completing a bachelors at Canisius who want to complete a Masters in Data Analytics is also available.

Data Analytics is a rapidly developing field driven by the need to effectively utilize Big Data to inform business decisions. It has the goal of making reliable predictions or inferences from very large collections of data drawn from a particular domain of human endeavor, including a wide range of diverse fields such as business management, science, sports, health-care management, criminal justice, and not-for-profit agencies.

The Master’s Program in Data Analytics at Canisius contains the three standard components of the field, namely, Statistics, Computer Science, and a Domain area, in Business. Eventually, we expect to expand the domain areas to include Insurance, the Health Sciences and other areas. Students with a strong interest in other domain areas are encouraged to discuss them with the Program Director. Domain courses are chosen to meet the students career goals and needs, and their prior academic and workplace experience.

In addition to offering the standard components of Data Analytics, Canisius will also focus on developing student capabilities in three crucial soft skills:

- the ability to work in collaborative, multidisciplinary teams;
- the ability to communicate effectively with different audiences, using a variety of written, oral, and visual modes of communication;
- a solid grounding in the ethics of data stewardship.

Consistent with the mission of the College, we have a strong interest in promoting collaborative efforts with the local non-profit agencies focused on the promotion of the public good. We will seek to connect interested students with analytic projects in conjunction with local not-for-profits.

While data analytics programs are rapidly being developed at many institutions, Canisius has a unique history with its focus on ethics, its emphasis on the ability to communicate with and understand others grounded in the Jesuit intellectual tradition, and the steadily increasing institutional emphasis on collaborative learning and teamwork. The tradition of high levels of personal attention to students at Canisius is the ideal environment for fostering these soft skills of communication, teamwork, and ethically grounded decision making, as well as the technical areas of computer programming and statistical inference.

Admissions Requirements

- Students from any undergraduate major are welcome to apply, as long as they have acquired a bachelor’s degree prior to the start of classes.
- Cumulative GPA of 2.8 or higher.
- Successful completion of a college-level Calculus 1 course (comparable to MAT 111 or MAT 115 at Canisius).
- Students may apply at any time. We have rolling admissions.
- Student preparation and background are used to determine if some introductory courses may be waived.

Materials to be Submitted

- Free Online Application (https://www.canisius.edu/admissions/apply-canisius/), with essay
- An official transcript from each college attended
- Official GRE or GMAT score (optional)
- Resume (optional)
- One or two Letters of Recommendations (optional)

Policies

Academic Standing
The Data Analytics program follows the College of Arts and Sciences policies on students’ academic standing, (http://catalog.canisius.edu/graduate/academics/academic-policies/#academicstandingtext)

Matriculation and Program Enrollment
The Data Analytics program follows the Canisius College policy for matriculated students (http://catalog.canisius.edu/graduate/admission-matriculation/#Matriculation) that expects students to maintain a continuous program of academic work.

Registration and Credit Hours
Data Analytics students must be registered for at least 4.5 credits per semester to maintain eligibility for financial aid (if they are eligible). A full load is at least 9 credit hours. No student may register for more than 12 credit hours in any semester.

Curriculum
This program is divided into three distinct components, comprising a total of at least 30 credit hours. The Preparatory Courses are base levels of knowledge and skill required before proceeding with the Core Competencies portion of the program. Up to 10 hours (3 courses) of the Preparatory Courses may be waived based on the student’s prior background and coursework. Students with exceptionally strong backgrounds may
substitute other domain courses (typically graduate business courses) for Preparatory courses. For example, this might occur for a student with an engineering degree, and thus strong computational and mathematical skills, or a finance degree with strong business and mathematical grounding.

Core Competencies portion consists of 5 courses, all of which were developed exclusively for the Data Analytics program. They cover advanced statistics, topics on managing data, as well as visualization/presentation.

The students will also participate in integrative projects in data analytics, gaining valuable hands-on experience and connections at companies in the Buffalo area and beyond.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Preparatory Courses (taken the year or summer prior to cohort start)</td>
<td></td>
</tr>
<tr>
<td>DAT 501</td>
<td>Statistics and Econometrics</td>
<td>3</td>
</tr>
<tr>
<td>CSC 511 &amp; 511L</td>
<td>Introduction to Programming and Introduction to Programming Lab</td>
<td>3</td>
</tr>
<tr>
<td>CSC 512 &amp; 512L</td>
<td>Data Structures and Algorithms and Data Structures and Algorithms Lab</td>
<td>3</td>
</tr>
<tr>
<td>MAT 500</td>
<td>Topics in Applied Mathematics</td>
<td>4</td>
</tr>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Elective (Domain specific)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Fall</td>
<td></td>
</tr>
<tr>
<td>DAT 511</td>
<td>Data Stewardship: Preparation, Exploration and Handling of Big Data</td>
<td>3</td>
</tr>
<tr>
<td>CSC 610 &amp; 610L</td>
<td>Database Management and Database Management Lab</td>
<td>3</td>
</tr>
<tr>
<td>DAT 521</td>
<td>Applied Integrative Projects in Data Analytics I</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Elective (Domain Specific)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Spring</td>
<td></td>
</tr>
<tr>
<td>DAT 512</td>
<td>Statistical Approaches to Big Data</td>
<td>3</td>
</tr>
<tr>
<td>DAT 514</td>
<td>Data Mining and Machine Learning</td>
<td>3</td>
</tr>
<tr>
<td>DAT 515</td>
<td>Visualization and Presentation of Advanced Analytics</td>
<td>3</td>
</tr>
<tr>
<td>DAT 522</td>
<td>Applied Integrative Projects in Data Analytics II</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Total Credits</td>
<td>40</td>
</tr>
</tbody>
</table>

1. Up to 10 credits of coursework (from those noted) may be waived by the program director based on a student’s preparation and experience.

## Domain Courses

Students will take at least two domain courses drawn from the courses below. Students may apply to the program director to take graduate level courses drawn from other domain areas, or more advanced courses for which they have adequate preparation.

### Business and Finance Domain

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACC 505</td>
<td>Financial Accounting</td>
<td></td>
</tr>
<tr>
<td>ECO 503</td>
<td>Statistics for Managers with Excel</td>
<td></td>
</tr>
<tr>
<td>FIN 608</td>
<td>Corporate Finance</td>
<td></td>
</tr>
<tr>
<td>FIN 617</td>
<td>Portfolio Analysis</td>
<td></td>
</tr>
<tr>
<td>FIN 619</td>
<td>Financial Modeling</td>
<td></td>
</tr>
<tr>
<td>FIN 620</td>
<td>Investment Management</td>
<td></td>
</tr>
</tbody>
</table>

### Investment Management

- Derivative Securities
- Fixed Income Securities

### Financial Modeling

- Statistical Approaches to Big Data
- Financial Accounting
- Statistics for Managers with Excel
- Corporate Finance
- Portfolio Analysis
- Financial Modeling
- Investment Management

### Portfolio Analysis

- Interactive Graphical Case Studies in Big Data
- Data Stewardship: Preparation, Exploration and Handling of Big Data
- Database Management

### Corporate Finance

- Applied Integrative Projects in Data Analytics I
- Applied Integrative Projects in Data Analytics II

### Roadmap

The following sequences are provided as examples, but students are strongly encouraged to work with the program director to determine the best sequence for the student’s background, experience, and interest.

#### Sample Progression, Full-time Study - Math Background

The following example is for a student with a mathematics degree (assuming at least one course each in statistics and computer programming). Note: this example results in waivers for MAT 500, DAT 501, and CSC 511L & CSC 511L

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td>CSC 512 &amp; 512L</td>
<td>Data Structures and Algorithms and Data Structures and Algorithms Lab</td>
<td>3</td>
</tr>
</tbody>
</table>

#### Sample Progression, Full-time Study - Computer Science Background

The following example is for a student with a computer science degree (assuming no statistics or advanced mathematics). Note: this example results in waivers for CSC 511L & CSC 511L and CSC 512 & CSC 512L.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td>DAT 501</td>
<td>Statistics and Econometrics</td>
<td>3</td>
</tr>
<tr>
<td>MAT 500</td>
<td>Topics in Applied Mathematics</td>
<td>4</td>
</tr>
</tbody>
</table>

#### Summer

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT 511</td>
<td>Data Stewardship: Preparation, Exploration and Handling of Big Data</td>
<td>3</td>
</tr>
<tr>
<td>CSC 610 &amp; 610L</td>
<td>Database Management and Database Management Lab</td>
<td>3</td>
</tr>
<tr>
<td>DAT 521</td>
<td>Applied Integrative Projects in Data Analytics I</td>
<td>2</td>
</tr>
</tbody>
</table>

#### Fall

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT 512</td>
<td>Statistical Approaches to Big Data</td>
<td>3</td>
</tr>
<tr>
<td>DAT 514</td>
<td>Data Mining and Machine Learning</td>
<td>3</td>
</tr>
<tr>
<td>DAT 515</td>
<td>Visualization and Presentation of Advanced Analytics</td>
<td>3</td>
</tr>
<tr>
<td>DAT 522</td>
<td>Applied Integrative Projects in Data Analytics II</td>
<td>3</td>
</tr>
</tbody>
</table>

#### Total Credits

- 30

---
DAT 514  Data Mining and Machine Learning  3
DAT 515  Visualization and Presentation of Advanced Analytics  3
DAT 522  Applied Integrative Projects in Data Analytics II  3

Total Credits  34

Sample Progression, Full-time Study - Business Background
The following example is for a student with a business degree (assuming a course in statistics or econometrics). Note: this example results in waivers for DAT 501 and 2 Domain Courses.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer (or prior to the start of the cohort)</td>
<td><strong>Code</strong></td>
<td><strong>Title</strong></td>
</tr>
<tr>
<td>CSC 511</td>
<td>Introduction to Programming</td>
<td>3</td>
</tr>
<tr>
<td>&amp; 511L</td>
<td>and Introduction to Programming Lab</td>
<td></td>
</tr>
<tr>
<td>CSC 512</td>
<td>Data Structures and Algorithms</td>
<td>3</td>
</tr>
<tr>
<td>&amp; 512L</td>
<td>and Data Structures and Algorithms Lab</td>
<td></td>
</tr>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td>MAT 500</td>
<td>Topics in Applied Mathematics</td>
<td>4</td>
</tr>
</tbody>
</table>

Fall
DAT 511  Data Stewardship: Preparation, Exploration and Handling of Big Data  3
CSC 610 & 610L Database Management and Database Management Lab  3
One Domain Course  3
DAT 521  Applied Integrative Projects in Data Analytics I  2

Spring
DAT 512  Statistical Approaches to Big Data  3
DAT 514  Data Mining and Machine Learning  3
DAT 515  Visualization and Presentation of Advanced Analytics  3
DAT 522  Applied Integrative Projects in Data Analytics II  3

Total Credits  34

Sample Progression, Full-time Study - STEM/Engineering Background
The following example is for a student with strong mathematics background (STEM/Engineering) including at least one course each in statistics and computer programming. Note: this example includes waivers for MAT 500, DAT 501, and CSC 511 & CSC 511L.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 1 (or prior to the start of the cohort)</td>
<td><strong>Code</strong></td>
<td><strong>Title</strong></td>
</tr>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td>CSC 512</td>
<td>Data Structures and Algorithms</td>
<td>3</td>
</tr>
<tr>
<td>&amp; 512L</td>
<td>and Data Structures and Algorithms Lab</td>
<td></td>
</tr>
</tbody>
</table>

Fall
DAT 521  Applied Integrative Projects in Data Analytics I  2
One Domain Course  3

Spring
CSC 610 & 610L Database Management and Database Management Lab  3
DAT 512  Statistical Approaches to Big Data  3

Summer 2
One Domain Course  3
Fall 2
DAT 511  Data Stewardship: Preparation, Exploration and Handling of Big Data  3
DAT 514  Data Mining and Machine Learning  3
Spring 2
DAT 522  Applied Integrative Projects in Data Analytics II  3
DAT 515  Visualization and Presentation of Advanced Analytics  3

Total Credits  30

Sample Progression, Part-time Study - Math Background
The following example is for a student with a mathematics degree (assuming at least one course each in statistics and computer programming). Note: this example results in waivers for MAT 500, DAT 501, and CSC 511 & CSC 511L.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer 1 (or prior to the start of the cohort)</td>
<td><strong>Code</strong></td>
<td><strong>Title</strong></td>
</tr>
<tr>
<td>CSC 511</td>
<td>Introduction to Programming</td>
<td>3</td>
</tr>
<tr>
<td>&amp; 511L</td>
<td>and Introduction to Programming Lab</td>
<td></td>
</tr>
</tbody>
</table>

Fall
DAT 511  Data Stewardship: Preparation, Exploration and Handling of Big Data  3
CSC 610 & 610L Database Management and Database Management Lab  3
One Domain Course  3
DAT 521  Applied Integrative Projects in Data Analytics I  2

Spring
DAT 512  Statistical Approaches to Big Data  3

Summer 2
One Domain Course  3
A solid grounding in the ethics of data stewardship, training that is highly relevant in an age where the accumulation and storage of personal data is ever increasing.

A total of 5 courses are required for the certificate. A minimum of 12 credits beyond those taken to fulfill the requirements for an undergraduate degree must be completed as part of the certificate.

**Code** | **Title** | **Credits**
---|---|---
MAT 500 | Topics in Applied Mathematics | 4
DAT 511 | Data Stewardship: Preparation, Exploration and Handling of Big Data | 3
DAT 514 | Data Mining and Machine Learning | 3
DAT 515 | Visualization and Presentation of Advanced Analytics | 3
DAT 522 | Applied Integrative Projects in Data Analytics II | 3

**Total Credits** | **34**

**Sample Progression, Part-time Study - Humanities, Social Sciences, or Education Background**

The following example is for a student with an undergraduate degree, but no programming background, or mathematics beyond a college Calculus 1 course.

**Code** | **Title** | **Credits**
---|---|---
DAT 500 | Interactive Graphical Case Studies in Big Data | 1
CSC 511 & 511L | Introduction to Programming and Introduction to Programming Lab | 3
CSC 512 & 512L | Data Structures and Algorithms and Data Structures and Algorithms Lab | 3
DAT 501 | Statistics and Econometrics | 3
DAT 521 | Applied Integrative Projects in Data Analytics I | 2
CSC 610 & 610L | Database Management and Database Management Lab | 3
One Domain Course | | 3
MAT 500 | Topics in Applied Mathematics | 4
One Domain Course | | 3
DAT 511 | Data Stewardship: Preparation, Exploration and Handling of Big Data | 3
DAT 514 | Data Mining and Machine Learning | 3
DAT 515 | Visualization and Presentation of Advanced Analytics | 3
DAT 522 | Applied Integrative Projects in Data Analytics II | 3
DAT 512 | Statistical Approaches to Big Data | 3

**Total Credits** | **34**

Note: Students without prior knowledge of the R programming language should take DAT 511.

Students who begin the program with the equivalent of MAT 500, will take an additional elective in lieu of MAT 500. These students may complete the certificate in 15 credits.

**Advanced Certificate in Data Analytics**

The Certificate in Data Analytics will enhance your ability work as a data specialist across a range of industries and organizations. This program is intended for professionals in related fields who want to utilize Data Analytics methods in their work, extending their existing skills but not completing a graduate degree. Students will learn how to work with and analyze large data sets, and how to apply data knowledge within a field of specialization. You’ll gain the effective teamwork and communication skills that are essential to working within diverse, multidisciplinary organizations. And you’ll get a solid grounding in the ethics of data stewardship, training that is highly relevant in an age where the accumulation and storage of personal data is ever increasing.

To satisfy DAT 501 Statistics and Econometrics, student may complete either ECO 455 or ECO 609.

MAT 500 Topics in Applied Mathematics can be waived for some majors, such as math, physics, or engineering. For other majors, MAT 211 or MAT 219 will NOT satisfy the MAT 500 requirement, but may help students prepare for the course.

Students may take CSC 111 and CSC 111L to satisfy CSC 511 and CSC 511L.

Students who have already completed CSC 111, may take CSC 112 and CSC 112L to satisfy the CSC 512 and CSC 512L requirement.

The following courses could be used to satisfy a domain course requirement: FIN 608, FIN 617, FIN 619, FIN 620, FIN 623, or FIN 628.
Learning Goals & Objectives
For the Master of Science degree

Student Learning Goal 1: Multi-disciplinary analytic capabilities.
- Objective A: Domain Knowledge: Students will be able to apply the computational and statistical methods and analytical tools to strategic and tactical decision making for at least one domain area. In business, for example, this might be: accounting, economics, finance, management, or marketing.
- Objective B: Adaptable grounding in applied statistics. Students will be able to use the basic principles of probability theory in a variety of contexts, including both classical statistical approaches and computational based methods. Students will be familiar with one modern statistical software platform and will be able to readily adapt to others.
- Objective C: Flexible computational skills. Students will have a strong working knowledge of at least one general purpose programming language, and will be able to work with a range of data structures within those languages. Students will also be familiar with databases and the programming techniques needed to work with Big Data.

Student Learning Goal 2: Effective teamwork.
- Objective A: Students will demonstrate the ability to work in multi-disciplinary teams to address real-world problems.
- Objective B: Students will understand the current theoretical ideas related to the formation of effective collaborative teams.

Student Learning Goal 3: Effective business communication.
- Objective A: Students will be able to identify the needs of different audiences, and effectively present complex information in ways that suit the needs of multiple audiences.
- Objective B: Students will be able to write effectively to convey data analytic results in business or other domain contexts.
- Objective C: Students will be able to create and deliver effective oral presentations, as well as present ideas in less formal oral settings.
- Objective D: Students will be able to create effective graphics, both static and real-time active displays, that convey results to business or other domain audiences.

Student Learning Goal 4: Ethical data stewardship.
- Objective A: Students will have an awareness of the ethic and moral issues that arise in working with large data sets, and understand the steps that need to be taken to protect the rights and privacy of the individuals involved.

For the Advanced Certificate
Student Learning Goal 1: To develop multi-disciplinary analytic capabilities of the students.
- Objective A: Students will be able to apply a broad range of computational and statistical methods, including both exploratory methods and predictive analytic tools to large data sets, using both local and distributed computer systems.
- Objective B: Students will possess a broad understanding of the principles of statistical reasoning, which will allow them to understand and assess the utility of new statistical tools as they become available, and to put those tools to practical use. Adoptable grounding in applied statistics.

- Objective C: Students will develop flexible computational skills. The students will be proficient in at least one general purpose programming language, and at least one modern statistical package. Students will be able to rapidly learn new languages and packages and maintain their own professional capabilities as new technology and procedures appear.

Student Learning Goal 2: To develop interdisciplinary teamwork skills and effective communication skills.
- Objective A: Students will be comfortable and effective working in multidisciplinary teams to address complex real-world problems.
- Objective B: Students will effectively communicate statistical and computational results, to a wide range of different audiences with different needs, using verbal, written, and visual modes of communication.

Student Learning Goal 3: To obtain knowledge and develop skills in ethical use of big data.
- Objective A: Students will have a strong grounding in the ethical use of big data.

Courses
Computer Science
CSC 511 Introduction to Programming 3 Credits
This foundational course will teach you the basics of computer programming using the Python language. You will design, code, test, and debug computer programs for textual and graphical applications.
Corequisite: CSC 511L.
Offered: every fall, spring, & summer.
CSC 511L Introduction to Programming Lab 0 Credits
Required lab for CSC 511.
Corequisite: CSC 511.
Offered: every fall, spring, & summer.
CSC 512 Data Structures and Algorithms 3 Credits
The primary focus of this course is data structures and their accompanying algorithms, including recursive algorithms. In order to judge between competing algorithms or alternative data structures, we will use analysis to discover the time and memory bounds of various approaches. We will also use object oriented programming as a useful way of constructing abstract data types and in general structuring complex programs. Several debugging tools and approaches will be explored, especially hand tracing of algorithms. The Python programming language will be our main vehicle.
Prerequisite: CSC 511 or CSC 111. Corequisite: CSC 512L.
Offered: every fall, spring, & summer.
CSC 512L Data Structures and Algorithms Lab 0 Credits
Required lab for CSC 512.
Corequisite: CSC 512.
Offered: every fall, spring, & summer.
CSC 530 Operating System Design and Distributed Computing 3 Credits
The design of operating system software, distributed applications, client/server and other models, security issues, and parallel programming on a High Performance Computing Cluster.
Prerequisite: A minimum grade of C in CSC 502 & CSC 502L. Corequisite: CSC 530L.
Offered: every fall.
CSC 530L Operating System Design and Distributed Computing Laboratory

Offered: once a year.

Required lab for CSC 530.
Prerequisite: A minimum grade of C in CSC 502 & CSC 502L. Corequisite: CSC 530L.
Offered: every fall.

CSC 610L Database Management Lab

Offered: every fall & spring.

CSC 610 Database Management

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.
Prerequisite: CSC 112 or CSC 512; may be taken concurrently.
Offered: every fall & spring.

CSC 610L Database Management Lab

Required lab for CSC 610.
Prerequisite: CSC 512L. Corequisite: CSC 610.
Offered: every fall & spring.

Data Analytics

DAT 111 Introduction to Reporting and Analysis

An introduction to the methods and tools for reporting quantitative data for decision support in a wide range of fields. This course is meant as an introductory course in the Data Science program, and for students in other disciplines preparing for decision support roles in a range of commercial, educational or research roles. Both the general theories and approaches to the presentation of data for decision support in tabular and graphic forms, and practical technical methods will be covered in the course. Most of the course time will be spent using Excel for these tasks, but Tableau and/or PowerBI as well as some basic SQL queries will also be covered. Whenever possible, “real-world” data drawn from a wide range of fields and disciplines will be used to illustrate problems and approaches to reporting of data.
Offered: every spring.

DAT 211 Advanced Statistics with R

This course is designed to introduce students to the programming language R. We will begin by talking about the benefits of R from a practical to an ethical level. Students will learn to install R and load packages. Students will then identify a data set they want to work with over the semester and preregister their project rationale, hypotheses, and analytic plan with OSF. Students will spend the majority of their time learning to execute their analytic plan in R. Students will present their project at Ignatian Scholarship Day. After their ISD presentation, students will archive their materials on OSF and update their preregistration to reflect any modifications made to the plan as they conducted their research, changes they would make if they were going to do the project again, and future analyses they would like to conduct with the data set.
Offered: once a year.

DAT 111 Econometrics

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. Throughout the course, we will use EXCEL, R, and SAS. We will cover most of Chapters 1-10 of the textbook, followed by some selected special topics as time permits. You should read through each chapter as we cover it. Special emphasis will be placed on conceptual understanding and application of econometric methods. For those who are interested in more involved discussions of the theoretical framework and/or the statistical or mathematical derivation behind any of the ideas discussed in class, feel free to meet with me outside of class.
Prerequisite: MAT 111 and CSC 111.
Offered: once a year.

DAT 412 Machine Learning

A foundational development of the core ideas and concepts in machine learning, with emphasis on the statistical foundations of machine learning but also applied work in Python, or a comparable language. Topics covered will include feature engineering and basis sets, gradient descent model fitting, kernel methods, model selection methods, bootstrapping and other permutation methods, model inference and averaging, tree based methods with boosting and bagging, neural nets and deep learning and graph based methods.
Prerequisite: MAT 219 and CSC 112.
Offered: once a year.

DAT 500 Interactive Graphical Case Studies in Big Data

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

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 seemingly harmless actions can pose threats to the information security of others.
Prerequisite: CSC 511 or CSC 111.
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, CSC 502 or equivalent, & MAT 500 or equivalent.
Offered: every spring.

DAT 513L Database Management Lab 0 Credits
Required lab for DAT 513.
Prerequisite: CSC 502L. Corequisite: DAT 513.
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, CSC 501, and CSC 502 or equivalents.
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 (courses may be taken concurrently) and the 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 is supervised internship or project course. Students may chose to apply for a competitive internship position in Data Analytics with a local corporation, government or not-for-profit agency, or may apply to carry out a data analytics project with an employer or on-campus research sponsor.
Prerequisites: DAT 500, DAT 514, DAT 521.
Offered: every fall, spring, & summer.

DAT 555 Seminar on Deep Learning 1 Credit
Deep Learning is a computational and mathematical approach to building "deep" or many layer neural network architectures for solving complex machine learning tasks, such as image processing, audio processing, complex time series, natural language processing and other big data problems. This course would teach students to build and training deep learning models using current state of the art tools.
Prerequisite: CSC 112 or CSC 512 and MAT 500 or MAT 211 or MAT 219.
Offered: occasionally.

Mathematics

MAT 500 Topics in Applied Mathematics 4 Credits
This course provides a brief overview of the basic tools from Linear Algebra and Multivariable calculus, with particular attention given to topics that are needed in Data Science. To facilitate students' understanding of the concepts, rigor and proofs will be de-emphasized while numerous examples will be discussed, including the use of computer software like MATLAB.
Prerequisites: One semester of Calculus (MAT 111 or MAT 115 at Canisius, or equivalent).
Offered: every summer.

MAT 591 Discrete Mathematics 3 Credits
Fundamental topics with computer science applications. Sets and logic, propositional and predicate calculus, elements of combinatorics and counting, elementary discrete probability, functions and relations, and graphs.
Offered: every fall.