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 (http://catalog.canisius.edu/mail: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. Khinks, 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
R. Mark Meyer, PhD, Associate Professor of Computer Science
Adina Oprisan, PhD, Assistant Professor of Mathematics and Statistics
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
Yuxing Paul Yan, MBA, PhD, Assistant Professor of Economics and Finance

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 particular, in Business or Health Care Information Technology. Eventually, we expect to expand the domain areas to include the Health Sciences and other areas. Students with a strong interest in other domain areas are encouraged to discuss them with the Program Director.

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)
- Resumé (optional)
- One or two Letters of Recommendations (optional)

Policies

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

Matriculation and Continued 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.

The following example is for a student with a computer science 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 501.

<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 502</td>
<td>Structures and Algorithms for Data Analytics and Introduction to Programming for Data Analytics 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>Elective (Domain specific)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Total Credits</td>
<td>30</td>
<td></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.

Sample Progression, Full-time Study - Math Background

The following example is for a student with a mathematics degree (assuming no statistics or advanced mathematics). Note: this example results in waivers for MAT 500, DAT 501, and CSC 501.

<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 502</td>
<td>Structures and Algorithms for Data Analytics and Introduction to Programming for Data Analytics 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>Elective (Domain specific)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Total Credits</td>
<td>30</td>
<td></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 501 and CSC 502.

<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>
<tr>
<td>One Domain Course</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

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.

<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>
</tbody>
</table>
### Fall
- DAT 511: Data Stewardship: Preparation, Exploration and Handling of Big Data
- DAT 513: Database Management
- One Domain Course
- DAT 521: Applied Integrative Projects in Data Analytics I

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

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></td>
<td></td>
</tr>
<tr>
<td>CSC 501</td>
<td>Introduction to Programming for Data Analytics</td>
<td>3</td>
</tr>
<tr>
<td>CSC 502</td>
<td>Structures and Algorithms for Data Analytics</td>
<td>3</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
- DAT 513: Database Management
- One Domain Course
- DAT 521: Applied Integrative Projects in Data Analytics I

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

Total Credits: 30

---

### 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 501.

<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></td>
<td></td>
</tr>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td>CSC 502</td>
<td>Structures and Algorithms for Data Analytics</td>
<td>3</td>
</tr>
</tbody>
</table>

### Fall
- DAT 511: Data Stewardship: Preparation, Exploration and Handling of Big Data
- DAT 513: Database Management
- One Domain Course
- DAT 521: Applied Integrative Projects in Data Analytics I

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

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 501.

<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></td>
<td></td>
</tr>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td>CSC 502</td>
<td>Structures and Algorithms for Data Analytics</td>
<td>3</td>
</tr>
</tbody>
</table>

### Fall
- DAT 511: Data Stewardship: Preparation, Exploration and Handling of Big Data
- DAT 513: Database Management
- One Domain Course
- DAT 521: Applied Integrative Projects in Data Analytics I

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

Total Credits: 30

---

### Sample Progression, Part-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 1 (or prior to the start of the cohort)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAT 500</td>
<td>Interactive Graphical Case Studies in Big Data</td>
<td>1</td>
</tr>
<tr>
<td>CSC 501</td>
<td>Introduction to Programming for Data Analytics</td>
<td>3</td>
</tr>
<tr>
<td>CSC 502</td>
<td>Structures and Algorithms for Data Analytics</td>
<td>3</td>
</tr>
</tbody>
</table>

### Fall
- DAT 511: Data Stewardship: Preparation, Exploration and Handling of Big Data
- DAT 513: Database Management
- One Domain Course
- DAT 521: Applied Integrative Projects in Data Analytics I

### Spring
- DAT 512: Statistical Approaches to Big Data
- DAT 514: Data Mining and Machine Learning
- DAT 515: Visualization and Presentation of Advanced Analytics
- MAT 500: Topics in Applied Mathematics

Total Credits: 30
The Certificate in Data Analytics will enhance your ability to 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.

### 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 I course.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fall 2</td>
<td></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>DAT 514</td>
<td>Data Mining and Machine Learning</td>
<td>3</td>
</tr>
<tr>
<td>Spring 2</td>
<td></td>
<td></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>Total Credits</td>
<td></td>
<td>34</td>
</tr>
</tbody>
</table>

### 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.

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.

<table>
<thead>
<tr>
<th>Code</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Required Courses</strong></td>
<td></td>
</tr>
<tr>
<td>MAT 500</td>
<td>Topics in Applied Mathematics</td>
<td>4</td>
</tr>
<tr>
<td>CSC 501 &amp; 501L</td>
<td>Introduction to Programming for Data Analytics and Introduction to Programming for Data Analytics Lab</td>
<td>3</td>
</tr>
<tr>
<td>CSC 502 &amp; 502L</td>
<td>Structures and Algorithms for Data Analytics and Structures and Algorithms for Data Analytics Lab</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td><strong>Elective Courses</strong></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>DAT 512</td>
<td>Statistical Approaches to Big Data</td>
<td>3</td>
</tr>
<tr>
<td>DAT 513</td>
<td>Database Management &amp; 513L and Database Management Lab</td>
<td>2</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 521</td>
<td>Applied Integrative Projects in Data Analytics I</td>
<td>2</td>
</tr>
<tr>
<td>Total Credits</td>
<td></td>
<td>16</td>
</tr>
</tbody>
</table>

**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.

### Preparation

Students interested in this degree will need specific skills to be successful. The following courses will help prepare you for the program, prior to the start of a formal cohort in the summer. We are providing the following advisement based on Canisius College courses. If you are considering taking courses elsewhere, please contact Michael Seaman (716-888-2545 or seamanm@canisius.edu) about courses at your institution that would provide appropriate backgrounds.

- To satisfy DAT 501 Statistics and Econometrics, student may complete either ECO 455 or ECO 526.
- 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 501.
- Students who have already completed CSC 111, may take CSC 112 and CSC 112L to satisfy the CSC 502 requirement.
- The following courses could be used to satisfy a domain course requirement: FIN 508, 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. Adaptable 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 501 Introduction to Programming for Data Analytics 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.
Offered: every fall, spring, & summer.

CSC 501L Introduction to Programming for Data Analytics Lab 0 Credits
Required lab for CSC 501.
Corequisite: CSC 501.
Offered: Summer.

CSC 502 Structures and Algorithms for Data Analytics 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 501 or CSC 111 as prerequisite.
Offered: every fall, spring, & summer.

CSC 502L Structures and Algorithms for Data Analytics Lab 0 Credits
Required lab for CSC 502.
Corequisite: CSC 502.
Offered: Summer.

Data Analytics

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.

Offered: every spring.

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.

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 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 513L Database Management Lab 0 Credits
Required lab for DAT 513.
Prerequisite: CSC 502L. Corequisite: DAT 513.
Offered: 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 511, CSC 512 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 will 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, S2-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.

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.