

THE ACADEMIC PROGRAM

New College offers a Master's Degree in Applied Data Science, under the Classification of Instructional Programs (CIP) code 11.0104. The degree requires 36 credit hours of graduate work. Students complete 11 courses at 3 credit hours each, one practicum at 3 credit hours, and non-credit earning courses (the Introduction to Data Science Boot Camp, the Industrial Workshops, a second industrial practicum, and the Industrial Seminar Series).

The Applied Data Science Degree Program emphasizes mathematical rigor and computational mastery in the collection, visualization, and use of data with particular focus on the statistical and computational challenges of very large and unstructured data sets. All New College graduate students enroll full-time; there are no part-time enrollment options for this program.

First Year

Pre-Semester

IDC 5100

First Semester

IDC 5204

IDC 5110

IDC 5120

IDC 5210

IDC 5251

January Interterm

IDC 5295

Second Semester

IDC 5205

IDC 5112

IDC 5130

IDC 5131

Year One Summer or Year Two January Interterm

IDC 6293

Second Year

Third Semester

IDC 6200

IDC 6215

IDC 6250

IDC 6253

Fourth Semester

IDC 6294

Graduate Certificates

New College offers four Graduate Certificate Programs integrated with its MS in Applied Data Science program. The purpose of these programs is to make its strong curriculum, and faculty expertise and experience, available to the community, including working professionals and anyone else who seeks to advance their knowledge and skills in the field of applied data science, analytics, and visualization for personal and professional development. Each graduate certificate program is offered as a 3-course sequence that participants will be taking for college credit. Each program is developed with a particular focus area in mind, such as data analytics and visualization, applied statistics, applied machine learning, and distributed computing, and is designed not only to provide

foundational knowledge but also to foster development of skills (e.g., programming, modeling) highly demanded in today's competitive job marketplace.

Admission Requirements

Participants must complete an online application for the intended certificate program and provide documentation of a Bachelor's degree (in any discipline).

Participants must also meet further requirements as specified by each certificate program (e.g. experience or knowledge in calculus, linear algebra and/or computer programming). Applicants must also submit a short letter of interest/intent to pursue their graduate certificate program of interest.

The graduate admissions committee will review each application and determine selections based on various factors including but not limited to: academic credentials from prior college attendance; knowledge, skills and experience in data science or a related field; interest in data science, applied statistics, business analytics, business intelligence systems or related domains; and overall motivation.

Completion Criteria

Participants of each graduate certificate program must achieve at least 3.0 cumulative GPA over all courses included in the program. A participant has the option of retaking a course (e.g., to increase their cumulative GPA to 3.0) regardless of their grade(s) in earlier attempt(s). After two or more attempts, only the most recent letter grade will be factored into the cumulative GPA calculation. Please note however that all course attempts will be reflected on final transcripts.

Upon successful completion of a graduate certificate program, the student will be provided by the Office of the Registrar a printed certificate and a transcript that details the coursework completed.

Graduate Certificate in Data Analysis and Visualization with R

This 3-course certificate program provides participants with a solid background in data analysis and data visualization using the widely used data science programming language R. Starting with a course that introduces the fundamental principles in data extraction, loading, pre-processing, and analysis, it continues with a course on applied statistics and concludes with a course on Data Visualization.

Prerequisites: None

Code	Title
IDC 5110	Data Munging and Exploratory Data Analysis
IDC 5204	Applied Statistics I
IDC 5112	Data Visualization

Learning Outcomes

- Demonstrate understanding of fundamental concepts in
 - Data cleansing and exploratory data analysis
 - Introductory statistical analysis
 - Data visualization and its role and importance in data analysis
- Construct effective data visualizations and communicate findings
- Demonstrate proficiency in R programming language
- Demonstrate awareness and recognition of ethical issues in data analysis and visualization

- Operate effectively in a teamwork environment and communicate effectively with peers
- Communicate orally and in writing with audiences the results of data analysis or visualization

Graduate Certificate in Machine Learning with Python

This 3-course certificate program aims to equip its participants with knowledge and skills to develop and apply machine learning models to solve complex real-world problems. Starting with a course that introduces the fundamental principles in algorithms and optimization, it continues with models in machine learning and concludes with a course on advanced topics in computing including neural networks and deep learning. All implementations are done in Python.

Prerequisites

- Knowledge of, or experience in, introductory Python
- Knowledge of linear algebra

Code	Title
IDC 5120	Algorithms for Data Science
IDC 5210	Applied Machine Learning
IDC 6215	Advanced Applied Computing

Learning Outcomes

- Demonstrate understanding of fundamental concepts in
 - Algorithms and optimization
 - Machine learning (ML) models and their applications
 - Advanced computational models, including deep learning (DL), and their applications
- Build data science pipelines that include ML/DL models
- Demonstrate proficiency in Python programming language
- Demonstrate awareness and recognition of ethical issues in machine learning and artificial intelligence
- Operate effectively in a teamwork environment and communicate effectively with peers
- Communicate orally and in writing with audiences the results of data analysis or visualization

Graduate Certificate in Statistical Modeling

This 3-course certificate program provides participants with a solid background in statistical modeling over a three-course sequence in statistics. Starting with fundamental concepts in descriptive and inferential statistics, it continues by exploring a variety of statistical models such as multivariate linear and logistic regression, time series modeling, survival analysis, Bayesian statistics, among others. Details regarding each course are provided below.

Prerequisites

- Knowledge of or experience in introductory R (unless the first course of **Graduate Certificate in Data Analysis and Visualization with R** is taken prior)

Code	Title
IDC 5204	Applied Statistics I
IDC 5205	Applied Statistics II
IDC 6200	Advanced Applied Statistics

Learning Outcomes

- Demonstrate understanding of fundamental concepts in
 - Descriptive and inferential statistics
 - Statistical modeling and computational techniques in statistical analysis
 - Various types of statistical models, including linear, logistic and generalized linear models, time series models, survival models and more
- Demonstrate proficiency in R programming language
- Demonstrate awareness and recognition of ethical issues in statistical modeling
- Operate effectively in a teamwork environment and communicate effectively with peers
- Communicate orally and in writing with audiences the results of data analysis or visualization

Graduate Certificate in Distributed Computing

This 3-course certificate program provides solid background as well as hands-on experience in distributed computing. Starting with a course that introduces the fundamental principles in algorithms and optimization, it continues with traditional and modern database systems including SQL and NoSQL databases. It concludes with a course on massively parallel datasets and database systems, and algorithms for parallel architectures. All implementations are done in Python.

Prerequisites

- Knowledge of, or experience in, introductory Python

Code	Title
IDC 5120	Algorithms for Data Science
IDC 5130	Databases for Data Science
	Distributed Computing

Learning Outcomes

- Demonstrate understanding of fundamental concepts in
 - Algorithms and optimization
 - Database systems; storage, retrieval and distribution of massive data sets
 - Parallel and distributed computing
- Demonstrate proficiency in Python programming language
- Operate effectively in a teamwork environment and communicate effectively with peers
- Communicate orally and in writing with audiences the results of data analysis or visualization