

DATA SCIENCE, MS

Degree: Master of Science (MS)

Field of Study: Data Science

Program Overview

MS in Data Science has three tracks: a Course-Focused track, a Project-Focused track, and a Thesis-Focused track. Although all of the three options require 30 credit hours, they are structured differently to achieve different objectives.

The Course-Focused track prepares students for advanced industry employment and should be treated as a terminal MS degree in Data Science. This track can be suitable for students who are coming from a non-computational background and looking to establish a solid background in data science.

The Project-Focused track is for students who seek opportunities for completing an applied project, for example within the context of an established collaboration with industry. This track can be suitable for students who come from a computational background and would like to specialize in an application field of data science or students who come from a non-computational background and would like to apply the data science knowledge they acquire to a project in their field.

The Thesis-Focused track is mainly for students who have interests in research. This track can also be suitable for students with computational background who would like to go into research and development in data science or students with a non-computational background who are looking to go into data-driven research in their field.

Therefore, the three tracks have different requirements in admission, advising, and course work.

Admission

Graduate students shall be admitted to the MS degree program upon recommendation of the faculty of the Department of Computer and Data Sciences. Requirements for admission include a strong record of scholarship in a completed bachelor's degree, and fluency in written and spoken English. The University requires all foreign applicants to show English proficiency by achieving a TOEFL score of at least 90 on the internet-based exam for the thesis-focused or the project-focused track.

For the course-focused track, a minimum TOEFL score of 80 is required.

For students who are expected to have any professional student to student interaction, e.g., as a teaching assistant, a lab instructor, or a tutor, a minimum TOEFL score of 90 is required. It is required that all students submit original copies of GRE scores, with the exception of CWRU students applying to the Combined Bachelor's/Master's Program.

The MS program requires students to have basic knowledge of computational programming and mathematical foundations of data science. The students who satisfy one of the following requirements are considered to have sufficient background knowledge for the Data Science MS program:

- Students who hold a Bachelor's degree in Computer Science or Data Science
- CWRU students who completed CWRU's minor in data science

c. Students who have completed the following course work:

- Programming in an object-oriented or functional language (e.g., Java, C++, Python) (e.g., CSDS 132)
- Data structures (e.g., CSDS 233)
- Linear algebra (e.g., MATH 201)
- Calculus-based statistics (e.g., STAT 312)

Applications from students who do not demonstrate sufficient knowledge in these fields may be granted admission on a provisional basis.

Students deficient in one or more of these areas (admission with provision) may be required to satisfy this requirement by taking the corresponding courses listed above. These courses cannot be counted towards their MS requirement. However, a student taking and passing a course that subsumes one of the requirements automatically demonstrates knowledge of the material in the required course; e.g. taking MATH 405 demonstrates knowledge of the material in MATH 201. Such graduate level courses may be used to satisfy their MS requirement.

Program Requirements

The Course-Focused MS degree program requirements consist of the completion of 30 credit hours of approved coursework, satisfactory completion of a comprehensive exam, i.e., passing the course ENGR 600 with a grade of "P". A student may pass ENGR 600 if they receive at least a B in at least two of the following courses:

Code	Title	Credit Hours
CSDS 413	Introduction to Data Analysis	3
CSDS 433	Database Systems	3
CSDS 435	Data Mining	3
CSDS 440	Machine Learning	3

The Project-Focused track requires 24 credit hours of coursework and 6 credit hours of CSDS 695.

The Thesis-Focused track requires 18 credit hours of coursework and 12 credit hours of CSDS 651. A BS/MS student is required to choose the thesis-focused track initially.

Both the Thesis-Focused and the Project-Focused track require a formal written report, as well as a final oral examination by a committee of at least three faculty members, two of whom must be primarily affiliated with the Department of Computer and Data Sciences. The academic advisor is normally one of the committee members. For Project-Focused track students, the oral examination fulfills the Comprehensive Examination requirement of the School of Graduate Studies.

If a student wishes to switch from one track to another, the following requirements apply:

- Deadline.** In each semester, students must request to switch track one week before the date at which Drop/Add ends, as stated in the academic calendar.
- Course-only or Project to Thesis.** A course-only student may request to switch to the thesis track only if she (1) has already taken at least 9 credit hours of letter graded CSDS courses and (2) has a GPA of 3.5

or higher and (3) has a TOEFL score of 90 or higher and (4) has the recommendation of a CDS advisor or (co)advisor.

- Course-only to Project. A course-only student may request to switch to the thesis track only if she (1) has a TOEFL score of 90 or higher and (2) has the recommendation of a CDS advisor or (co)advisor.
- Thesis to Project, or Thesis or Project to Course-only. Such a transfer needs approval from the student's advisor and the department chair.
- Petition. If a student fails to satisfy the transfer requirements, a petition may be submitted by a CDS advisor or (co)advisor to the department chair. In no case, petitions may be submitted by non-CDS faculty members or by students.

Students should consult with their academic advisor and/or department to determine the detailed requirements within this framework.

Course Requirements

For all three tracks, at least 24 credit hours of coursework must be at the 400-level or above. The course requirements for the Data Science MS are as follows:

- Students must pass CSDS 413
- Students must pass at least three courses from the courses listed under "Computational Foundations", "Analytics & Intelligence", and "Data Science in the Field" below. At least two of these courses must be from two different lists and the third course can be from any of these three lists.
- The remaining courses can be any course from these three lists, courses from the list titled "Suggested and Pre-approved Electives" below, or any course that is approved by the graduate affairs committee.
- At most three of the courses used to satisfy the degree requirements can be non-CSDS courses.

The courses are categorized as follows:

Code	Title	Credit Hours
Computational Foundations:		
CSDS 410	Analysis of Algorithms	3
CSDS 433	Database Systems	3
CSDS 438	High Performance Data and Computing	3
CSDS 444	Computer Security	3
CSDS 456	Data Privacy	3
CSDS 477	Advanced Algorithms	3
CSDS 494	Introduction to Information Theory	3

Code	Title	Credit Hours
Analytics and Intelligence		
CSDS 435	Data Mining	3
CSDS 440	Machine Learning	3
CSDS 442	Causal Learning from Data	3
CSDS 491	Artificial Intelligence: Probabilistic Graphical Models	3

CSDS 496	Artificial Intelligence: Sequential Decision Making	3
CSDS 497	Artificial Intelligence: Statistical Natural Language Processing	3

Code	Title	Credit Hours
------	-------	--------------

Data Science in the Field

CSDS 458	Introduction to Bioinformatics	3
CSDS 459	Bioinformatics for Systems Biology	3
CSDS 478	Computational Neuroscience	3
CSDS 461	Biomedical Image Processing and Analysis	3
DSCI 430	Cognition and Computation	3
DSCI 452	Applied Data Science Research	3
BUAI 411	Operations Analytics: Deterministic	3
BUAI 432	Operations Analytics: Stochastic	3
BUAI 435	Marketing Models and Digital Analytics	3
BUAI 446	Machine Learning and Artificial Intelligence in Business Analytics	3

Code	Title	Credit Hours
------	-------	--------------

Suggested and Pre-Approved Electives

STAT 425	Data Analysis and Linear Models	3
STAT 426	Multivariate Analysis and Data Mining	3
STAT 433	Uncertainty in Engineering and Science	3
STAT 448	Bayesian Theory with Applications	3
STAT 455	Linear Models	3
MATH 427	Convexity and Optimization	3
MATH 431	Introduction to Numerical Analysis I	3
MATH 439	Bayesian Scientific Computing	3
MATH 444	Mathematics of Data Mining and Pattern Recognition	3
DSCI 451	Exploratory Data Science	3
DSCI 453	Data Science: Statistical Learning, Modeling and Prediction	3
DSCI 454	Data Visualization and Analytics	3
PQHS 430	Basics of Probability and Statistical Theory	0
PQHS 431	Statistical Methods I	3
PQHS 432	Statistical Methods II	3
PQHS 453	Categorical Data Analysis	3
PQHS 459	Longitudinal Data Analysis	3
PQHS 471	Machine Learning & Data Mining	3
PQHS 480	Introduction to Mathematical Statistics	3
PQHS 550	Meta-Analysis & Evidence Synthesis	2 - 3

The students must have a Grade Point Average (GPA) of at least 3.00/4.00 to receive their M.S. degree.