

Data Science Secondary Field

Graduate students across Harvard can complete a secondary field in Data Science. This secondary field is available to any student enrolled in a PhD program in the Graduate School of Arts and Sciences upon approval of a plan of study by the Data Science Program Committee and the director of graduate studies in the student's home department.

Data Science lies at the intersection of statistical methodology, computational science, and a wide range of application domains. This Secondary Field offers strong preparation in statistical modeling, machine learning, optimization, management and analysis of massive data sets, and data acquisition. Students completing the Data Science Secondary Field will be exposed to topics such as reproducible data analysis, collaborative problem solving, visualization and communication, and security and ethical issues that arise in data science.

The Data Science Secondary Field is overseen by the joint leadership of the Computer Science and Statistics faculties and administered by the Institute for Applied Computational Science (IACS). All questions should be directed to Daniel Weinstock, Associate Director of Graduate Studies (ADGS) in Applied Computation (dweinsto@seas.harvard.edu)

Admission

Admission into the Data Science secondary field is by application, which must be submitted to the ADGS in Applied Computation (Maxwell Dworkin G107). Students interested in the secondary field should consult with their departmental director of graduate studies no later than the first semester of the third year of study. Applications may be submitted twice a year, in the spring semester (deadline: March 1) and fall semester (deadline: October 1) for the following academic term. The application, which will include a proposed Plan of Study, must also be approved by the home department DGS. The ADGS in Applied Computation will respond to all applications within one month.

Requirements

Each student's plan of study for the secondary field will include:

1. Core Courses

At least 3 of the Data Science core courses:

- AC 209a* Data Science 1: Introduction to Data Science
- AC 209b* Data Science 2: Advanced Topics in Data Science
- AM 207 Advanced Scientific Computing: Stochastic Methods for Data Analysis, Inference, and Optimization
- CS 207 Systems Development for Computational Science
- AC 221 Critical Thinking in Data Science

*Students can, with the permission of the program committee, count CS 109a/b in place of AC 209a/b.

2. Electives

Two electives in Computer Science or Statistics. Students may choose from a wide variety of elective courses offered by the Computer Science and Statistics faculties. A list of suggested electives can be found here:

<https://www.seas.harvard.edu/programs/graduate/applied-computation/master-of-science-in-data-science/data-science-courses>

Alternatively, students may choose to satisfy the elective requirement by taking additional core courses. Students may also choose, as a substitute for one elective, either AC 297r, the IACS Capstone Project course, or AC298r, the Interdisciplinary seminar in Computational and Data Science.

3. Oral Examination

As a final requirement, an oral examination by a faculty committee on a data science research topic. Typically students will present on a part of their dissertation thesis work. Students will be evaluated on their ability to explain their work to the interdisciplinary IACS audience and their command of the Data Science methods they have used. The oral presentation should explain how the courses taken to satisfy the Data Science Secondary Field impact their research.

Advising and Other Activities

Daniel Weinstock, ADGS in Applied Computation, will be responsible for front-line advising of students in the Data Science Secondary Field. Students interested in the secondary field are encouraged to reach out to Dr. Weinstock before submitting an application. Students enrolled in the Secondary Field will be able to participate in the activities of the IACS community, including technical and interdisciplinary colloquia, skill-building workshops, and tech-treks to local companies working to apply computation and data science in many different domains.