Data science is the science of extracting actionable knowledge from large and complex data repositories, where “complex” may refer to the modality of the data (images, time series, text, as well as traditional tabular data) or other facets of the data in question (data can be complex because they are geographically distributed, or characterized by the ubiquity of missing or inaccurate values). As such, data science is an interdisciplinary field of study, encompassing sub-areas of computer science, statistics, electrical engineering, and applied mathematics. Furthermore, data science has quickly become a critical enabling capability in many different fields—science, healthcare, energy, manufacturing, and many others. In all of these fields, the amount of data that can be usefully collected is growing exponentially. The next breakthroughs in these fields are likely to come from those who can best analyze the large amounts of data they are able to collect and convert that analysis into actionable insights.

At the undergraduate-level, the minor in Data Science is administered by the Data to Knowledge Lab (D2K).

At the graduate level, the Master of Data Science (MDS) degree is a professional degree administered by the Computer Science department for students intending to pursue a technical career. There is an on-premise and a fully online option. Students are admitted directly into one or the other option and cannot switch between the two, but the resulting degree is the same.

Minor

- Minor in Data Science [link]

Master’s Programs

- Master of Data Science (MDS) Degree [link]
- Master of Data Science (MDS) Degree, Online Program [link]

Director, Data to Knowledge Lab (D2K)

Xia (Ben) Hu, Computer Science

Program Directors

Su Chen (D2K Lab and DSCI Minor), Electrical and Computer Engineering
DSCI 301 - PROBABILITY AND STATISTICS FOR DATA SCIENCE

Short Title: STATISTICS FOR DATA SCIENCE

Department: Data Science

Grade Mode: Standard Letter

Course Type: Lecture/Laboratory

Distribution Group: Distribution Group III

Credit Hours: 4

Restrictions: Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.

Course Level: Undergraduate Upper-Level

Prerequisite(s): MATH 102 or MATH 106

Description: An introduction to mathematical statistics and computation for applications to data science. Topics include probability, random variables, expectation, sampling distributions, estimation, confidence intervals, hypothesis testing, and regression. A weekly lab will cover the statistical package, R, and data projects. Cross-list: STAT 315.

Recommended Prerequisite(s): MATH 212. Mutually Exclusive: Cannot register for DSCI 301 if student has credit for BUSI 395.

DSCI 302 - INTRODUCTION TO DATA SCIENCE TOOLS AND MODELS

Short Title: DATA SCIENCE TOOLS AND MODELS

Department: Data Science

Grade Mode: Standard Letter

Course Type: Lecture

Credit Hours: 3

Restrictions: Enrollment is limited to students with a minor in Data Science. Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.

Course Level: Undergraduate Upper-Level

Prerequisite(s): COMP 140 or DSCI 101

Description: This course introduces key concepts in data management, preparation, and modeling and provides students with hands-on experience in performing these tasks using modern tools, including relational databases, pandas, and Spark. Models covered include k-Nearest Neighbors, linear regression, and gradient descent. For registration purposes, COMP 140 is a required prerequisite for this course. With instructor permission, students who have experience with the Python programming language may be allowed to special register for this course. Note that these students may be required to demonstrate proficiency with Python. Priority for this course is given to students enrolled in the data science minor. Other students may be permitted to enroll at the discretion of the instructor.

DSCI 303 - MACHINE LEARNING FOR DATA SCIENCE

Short Title: MACHINE LEARNING FOR DS

Department: Data Science

Grade Mode: Standard Letter

Course Type: Lecture

Credit Hours: 3

Restrictions: Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.

Course Level: Undergraduate Upper-Level

Prerequisite(s): (DSCI 101 or COMP 140) and DSCI 302 and (DSCI 301 or STAT 310 or STAT 280 or STAT 305 or ELEC 303 or PSYC 339 or SOCI 382 or BIOE 439 or ECON 307)

Description: This course is a practical introduction to machine learning, emphasizing when and how to apply techniques and how to interpret results. Topics covered include regression, classification, dimension reduction, clustering, decision trees, ensemble learning, and neural networks. Mutually Exclusive: Cannot register for DSCI 303 if student has credit for ELEC 478/ELEC 578.

DSCI 304 - INTRODUCTION TO EFFECTIVE DATA VISUALIZATION

Short Title: DATA VISUALIZATION

Department: Data Science

Grade Mode: Standard Letter

Course Type: Lecture/Laboratory

Credit Hours: 3

Restrictions: Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.

Course Level: Undergraduate Upper-Level

Prerequisite(s): (DSCI 301 or ECON 307 or STAT 310 or STAT 315) and (DSCI 302 (may be taken concurrently) or COMP 330 (may be taken concurrently))

Description: This course teaches fundamental data visualization skills to undergraduate students in the Data Science minor. Students will learn how to create data visualizations in Python or R, how to design effective visualizations that account for visual perception, and how to explain and present data to technical and non-technical audiences.

DSCI 305 - DATA, ETHICS, AND SOCIETY

Short Title: DATA, ETHICS, AND SOCIETY

Department: Data Science

Grade Mode: Standard Letter

Course Type: Seminar

Distribution Group: Distribution Group II

Credit Hours: 3

Restrictions: Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.

Course Level: Undergraduate Upper-Level

Description: An examination of the ethical implications and societal impacts of choices made by data science professionals. The course will provide practical guidance on evaluating ethical concerns, identifying the potential for harm, and applying best practices to protect privacy, design responsible algorithms, and increase the societal benefit of data science research.

DSCI 400 - DATA SCIENCE AND MACHINE LEARNING SELF-GUIDED CAPSTONE LABORATORY

Short Title: DATA SCIENCE CAPSTONE LAB

Department: Data Science

Grade Mode: Standard Letter

Course Type: Lecture/Laboratory

Credit Hours: 3

Restrictions: Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.

Course Level: Undergraduate Upper-Level

Prerequisite(s): (DSCI 301 or STAT 315 or STAT 310 or ECON 307) and (DSCI 302 or COMP 330) and (DSCI 303 or STAT 413 or COMP 540) and DSCI 304

Description: In this project-based course, student teams will choose, define, and execute semester-long data-science and machine-learning research projects. These projects may be selected from a variety of disciplines and industries, where freedom is given in defining the projects. The course is about learning best practices in data science and machine learning while finding a suitable curiosity-driven project to build these methods and systems around.
Description and Code Legend

Note: Internally, the university uses the following descriptions, codes, and abbreviations for this academic program. The following is a quick reference:

Course Catalog/Schedule
- Course offerings/subject code: DSCI

Department (or Program) Description and Code
- Data Science: DSCI

Undergraduate Minor Description and Code
- Minor in Data Science: DSCI

Graduate Degree Description and Code
- Master of Data Science degree: MDS

Graduate Degree Program Description and Code
- Degree Program in Data Science: DATA

Graduate Degree Program Option Description and Code*
- Degree Program Option - Online (MDS degree only): OMDS

CIP Code and Description

1. Systems Use Only: this information is used solely by internal offices at Rice University (such as OTR, GPS, etc.) and primarily within student information systems and support.

CIP Code and Description 1
- DATA Major/Program: CIP Code/Title: 30.7001 - Data Science, General
- DSCI Minor: CIP Code/Title: 30.7001 - Data Science, General

Classification of Instructional Programs (CIP) 2020 Codes and Descriptions from the National Center for Education Statistics: https://nces.ed.gov/ipeds/cipcode/ (https://nces.ed.gov/ipeds/cipcode/)

2022-2023 General Announcements PDF Generated 10/31/22