DATA SCIENCE

Contact Information

Data Science
https://datascience.rice.edu

Su Chen
Assistant Teaching Professor, D2K Lab; Director, DSCI Minor
su.chen@rice.edu

Christopher M. Jermaine
Chair, Computer Science; Director, MDS degree program
christopher.m.jermaine@rice.edu

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 (https://ga.rice.edu/programs-study/departments-programs/engineering/data-science/data-science-minor/)

Master's Programs

- Master of Data Science (MDS) Degree (https://ga.rice.edu/programs-study/departments-programs/engineering/data-science/data-science-mds/)
- Master of Data Science (MDS) Degree, Online Program (https://ga.rice.edu/programs-study/departments-programs/engineering/data-science/data-science-mds-online/)

Director, Data to Knowledge Lab (D2K)

Xia (Ben) Hu, *Computer Science*

Program Directors

Su Chen (D2K Lab and DSCI Minor), *Electrical and Computer Engineering*

Christopher M. Jermaine (MDS Degree Program), *Computer Science*

Minor Advisors

Arko Barman, *Data to Knowledge Lab, Electrical and Computer Engineering*
Su Chen, *Data to Knowledge Lab, Electrical and Computer Engineering*
Hua Gong, *Sport Management*
Rudy Guerra, *Statistics*
Christopher M. Jermaine, *Computer Science*
Elizabeth Petrick, *History*

Steering Committee

David Alexander, *Physics and Astronomy*
Rudy Guerra, *Statistics*
Matthias Heinkenschloss, *Computational and Applied Mathematics*
Christopher M. Jermaine, *Computer Science*
Luay K. Nakhleh, *Computer Science*
Barbara Ostdiek, *Finance and Statistics*
Kirsten Ostherr, *English*
Frederick L. Oswald, *Psychological Sciences*
Renata Ramos, *Bioengineering*
Devika Subramanian, *Computer Science*
Marina Vannucci, *Statistics*
Ashok Veeraraghavan, *Electrical and Computer Engineering*
Jennifer Wilson, *Program in Writing and Communication*

For Rice University degree-granting programs:
To view the list of official course offerings, please see Rice's Course Catalog (https://courses.rice.edu/admweb/!SWKSCAT.cat?p_action=cata)
To view the most recent semester's course schedule, please see Rice's Course Schedule (https://courses.rice.edu/admweb/!SWKSCAT.cat)

DSCI 101 - INTRODUCTION TO DATA SCIENCE
Short Title: INTRO TO DATA SCIENCE
Department: Data Science
Grade Mode: Standard Letter
Course Type: Lecture/Laboratory
Distribution Group: Distribution Group III
Credit Hours: 3
Restrictions: Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.
Course Level: Undergraduate Lower-Level
Description: In this course, students learn the fundamentals of data science and Python programming while working on teams to solve real data science challenges, design a data science pipeline, and derive and communicate valuable insights from data. This is a non-calculus based course with no prior background in statistics or programming required.

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 ECON 307/STAT 310.

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 major in Sport Analytics. 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 kNearest Neighbors, linear regression and gradient descent. For registration purposes, DSCI 101 or 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 or sport analytics major. Other students may be permitted to enroll at the discretion of the instructor. Mutually Exclusive: Cannot register for DSCI 302 if student has credit for COMP 330.

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 (may be taken concurrently) or COMP 330 or COMP 430) and (DSCI 301 or STAT 310 or STAT 280 or STAT 305 or ELEC 303 or PSYC 339 or SOCI 382 or SOSC 302 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.

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): BIOE 439 or DSCI 301 or ECON 307 or ECON 310 or COMP 330 or COMP 340 or STAT 280 or STAT 305 or STAT 311 or STAT 312 or STAT 310 or STAT 315 or SOCI 382 or SOSC 302 or BUSI 251 or BUSI 395 or ELEC 303
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 STAT 311 or ECON 307 or ELEC 303 or BIOE 439 or SOCI 382 or PSYC 339) and (DSCI 302 or COMP 330 or COMP 430) and (DSCI 303 or ELEC 478 or STAT 413 or COMP 540)
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.

DSCI 415 - DATA SCIENCE CONSULTING
Short Title: DATA SCIENCE CONSULTING
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): STAT 405 or COMP 140 or CAAM 210
Description: Students in this course will advise clients at Rice and beyond in a data science consulting clinic, learn best practices in consulting, and gain exposure to a variety of real data science problems. Graduate/Undergraduate Equivalency: DSCI 515. Mutually Exclusive: Cannot register for DSCI 415 if student has credit for DSCI 515. Repeatable for Credit.

DSCI 435 - APPLIED MACHINE LEARNING AND DATA SCIENCE PROJECTS
Short Title: DATA SCIENCE PROJECTS
Department: Data Science
Grade Mode: Standard Letter
Course Type: Lecture/Laboratory
Credit Hours: 4
Restrictions: Enrollment is limited to Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.
Course Level: Undergraduate Upper-Level
Description: In this project-based course, student teams will complete semester-long data science research or analysis projects selected from a variety of disciplines and industries. Students will also learn best practices in data science. Cross-list: COMP 449. Graduate/Undergraduate Equivalency: DSCI 535. Repeatable for Credit.

DSCI 515 - DATA SCIENCE CONSULTING
Short Title: DATA SCIENCE CONSULTING
Department: Data Science
Grade Mode: Standard Letter
Course Type: Lecture/Laboratory
Credit Hours: 3
Restrictions: Enrollment is limited to Graduate level students.
Course Level: Graduate
Description: Students in this course will advise clients from across this Rice community in a data science consulting clinic, learn best practices in consulting, and gain exposure to a variety of real data science problems. Graduate/Undergraduate Equivalency: DSCI 415. Mutually Exclusive: Cannot register for DSCI 515 if student has credit for DSCI 415. Repeatable for Credit.

DSCI 535 - APPLIED MACHINE LEARNING AND DATA SCIENCE PROJECTS
Short Title: DATA SCIENCE PROJECTS
Department: Data Science
Grade Mode: Standard Letter
Course Type: Lecture/Laboratory
Credit Hours: 4
Restrictions: Enrollment limited to students in the MDS, OMCS or OMDS programs. Enrollment is limited to Graduate level students.
Course Level: Graduate
Description: In this project-based course, student teams will complete semester-long data science research or analysis projects selected from a variety of disciplines and industries. Students will also learn best practices in data science. Cross-list: COMP 549. Graduate/Undergraduate Equivalency: DSCI 435. Repeatable for Credit.

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]
- **DATA** Major/Program: CIP Code/Title: *30.7001 - Data Science, General*
- **DSCI** Minor: CIP Code/Title: *30.7001 - Data Science, General*

*   *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.*

[1]   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/)