

DATA SCIENCE

Contact Information

Data Science

<https://datascience.rice.edu>

Christopher M. Jermaine

Program Director

christopher.m.jermaine@rice.edu

Rudy Guerra

Minor Advisor

rguerra@rice.edu

The minor in Data Science is an interdisciplinary undergraduate program administered by the George R. Brown School of Engineering.

The Data Science minor curriculum emphasizes doing data science and aims to teach students best practices in the field. Students learn technical competencies by taking core courses in statistics, computer science, and machine learning. This knowledge base is complemented with courses that inform the student of the broader impact of the information age on human activity, including discussions on data privacy, ethics, reproducibility, communication, decision-making, and data visualization. This program culminates with a capstone experience whereby students work in teams to complete a semester-long data science project selected from a variety of disciplines and industries. The curriculum is summarized in terms of four foundational competencies: quantitative, communications, ethics, and substantive application.

Minor

- [Minor in Data Science \(https://ga.rice.edu/programs-study/departments-programs/engineering/data-science/data-science-minor/\)](https://ga.rice.edu/programs-study/departments-programs/engineering/data-science/data-science-minor/)

Data Science does not currently offer an academic program at the graduate level.

Program Director

Christopher M. Jermaine, *Computer Science*

Minor Advisors

Arko Barman, *Electrical and Computer Engineering, and Data to Knowledge Lab*

Su Chen, *Data to Knowledge Lab*

Rudy Guerra, *Statistics*

Christopher M. Jermaine, *Computer Science*

Frederick L. Oswald, *Psychological Sciences*

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, Biochemistry and Cell Biology*

Barbara Ostdiek, *Finance and Statistics*

Kirsten Osther, *English*

Frederick L. Oswald, *Psychological Sciences*

Renata Ramos, *Bioengineering*

Devika Subramanian, *Computer Science, Electrical and Computer Engineering*

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 \(https://courses.rice.edu/admweb!/SWKSCAT.cat?p_action=cata\)](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 \(https://courses.rice.edu/admweb!/SWKSCAT.cat\)](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

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 or MATH 112

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 Undergraduate, Undergraduate Professional or Visiting Undergraduate level students.**Course Level:** Undergraduate Upper-Level**Prerequisite(s):** COMP 140 and (DSCI 301 or ECON 307 or STAT 310 or STAT 315)**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 and Spark. Models covered include linear and logistic regression and gradient descent. For registration purposes, COMP 140 is a required prerequisite for this course. With instructor permission, students that have taken CAAM 210 (or another applicable course) may be allowed to special register for this course. Students seeking this instructor permission (to waive or substitute the COMP 140 prerequisite requirement) are expected to know the Python programming language, and may be required to demonstrate proficiency.**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 301 or STAT 315 or STAT 310) and (DSCI 302 or COMP 330)**Description:** This course is an introduction to concepts, methods, best practices, and theoretical foundations of machine learning. Topics covered include regression, classification, kernels, dimensionality reduction, clustering, decision trees, ensemble learning, regularization, learning theory, and neural networks. Recommended Prerequisite(s): CAAM 334 or CAAM 335 or MATH 355 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.**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 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

Program Description and Code

- Data Science: DSCI

Undergraduate Minor Description and Code

- Minor in Data Science: DSCI

CIP Code and Description ¹

- **DSCI Minor:** CIP Code/Title: 27.0304 - Computational and Applied Mathematics

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/>)