Data Science, BS

Learning Outcomes

Data Curation Skills
Graduates will be able to:

• understand issues associated with data collection, management, provenance, storage, merging, sharing, and preparation;
• work with multiple-source, multiple-format data;
• investigate the quality of the data; and
• have a basic understanding of ethical and confidentiality issues associated with data collection, storage, merging, and sharing.

Computational Skills
Graduates will be able to:

• use critical thinking skills to translate substantive questions into well-defined computational problems and choose appropriate computational techniques for a given problem;
• understand the foundational software skills and associated algorithmic and computational problem-solving methods used in computer science;
• be proficient in computational methods for collecting, managing, preparing, sharing, and describing data numerically and graphically from a variety of sources to design and carry out basic simulation studies; and
• use professional statistical software and understand the principles of programming and algorithmic problem solving that underlie these packages.

Statistical/Probabilistic Skills
Graduates will be able to:

• use critical thinking skills to translate substantive questions into well-defined statistical or probability problems and choose the appropriate graphical or numerical descriptive and/or inferential statistical techniques for a given problem;
• understand the importance of, and issues related to, the choice of the study design, such as designed experiment versus probability sample versus convenience sample, used to produce data;
• understand that uncertainty, variability, and randomness play significant roles in data-driven decision-making;
• understand how to measure and display uncertainty, the effect of randomness, confidence/credibility, and the likelihood of incorrect inferences;
• understand and be able to explain common misperceptions, paradoxes, and fallacies of probability and statistics; and
• understand basic regression, prediction, simulation, and visualization methods.

Mathematical Skills
Graduates will:

• have a firm grasp of the mathematical tools underlying statistical and computational methods which are primarily based on ideas in calculus, linear algebra, and discrete mathematics, including distribution theory, uncertainty quantification (e.g., probability theory), the probabilistic basis of formal statistical inference, models, and algorithms, and combinatorial analysis and recursion, which are used for algorithmic analysis, design, and for distribution theory.

Communication Skills
Graduates will be able to:

• clearly justify and communicate study results to a nontechnical audience;
• write accurate and meaningful reports that describe the statistical and computational analyses and summarize important findings; and
• work effectively as part of a team to address substantive questions that can be handled using statistical and computational methods.

Requirements

The Bachelor of Science with a major in data science requires a minimum of 120 s.h., including at least 59 s.h. of work for the major. Students must maintain a grade-point average of at least 2.00 in all courses for the major and in all UI courses for the major. They also must complete the College of Liberal Arts and Sciences GE CLAS Core.

Data science majors may not earn a major or minor in computer science or statistics, a major in computer science and engineering, or the Certificate in Social Science Analytics.

The BS with a major in data science requires the following coursework.

Prerequisite Courses

<table>
<thead>
<tr>
<th>Course #</th>
<th>Title</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>These:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATH:1550</td>
<td>Engineering Mathematics I: Single Variable Calculus</td>
<td>4</td>
</tr>
<tr>
<td>MATH:1560</td>
<td>Engineering Mathematics II: Multivariable Calculus</td>
<td>4</td>
</tr>
<tr>
<td>MATH:2700</td>
<td>Introduction to Linear Algebra</td>
<td>4</td>
</tr>
<tr>
<td>Or these:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MATH:1850</td>
<td>Calculus I</td>
<td>4</td>
</tr>
<tr>
<td>MATH:1860</td>
<td>Calculus II</td>
<td>4</td>
</tr>
<tr>
<td>MATH:2700</td>
<td>Introduction to Linear Algebra</td>
<td>4</td>
</tr>
<tr>
<td>MATH:2850</td>
<td>Calculus III</td>
<td>4</td>
</tr>
</tbody>
</table>

Core Courses

<table>
<thead>
<tr>
<th>Course #</th>
<th>Title</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>All of these:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS:1210</td>
<td>Computer Science I: Fundamentals</td>
<td>4</td>
</tr>
<tr>
<td>CS:2210</td>
<td>Discrete Structures</td>
<td>3</td>
</tr>
<tr>
<td>CS:2230</td>
<td>Computer Science II: Data Structures</td>
<td>4</td>
</tr>
</tbody>
</table>
CS:3330  Algorithms  3
STAT:2010  Statistical Methods and Computing  3
STAT:3100/IGPI:3100  Introduction to Mathematical Statistics I  3
STAT:3101/IGPI:3101  Introduction to Mathematical Statistics II  3

Advanced Courses

Course #  Title  Hours
Both of these:
CS:4400  Database Systems  3
STAT:4580/DATA:4580/IGPI:4580  Data Visualization and Data Technologies  3
One of these:
CS:5430  Machine Learning  3
STAT:4540/BAIS:4540/DATA:4540/IGPI:4540  Statistical Learning  3

Advanced Electives

Course #  Title  Hours
Three of these, with at least one computer science course and one statistics course:
DATA:4750  Probabilistic Statistical Learning  3
ACTS:6200/DATA:6200/STAT:6200  Predictive Analytics  3
CS:4420  Artificial Intelligence  3
CS:4440  Web Mining  3
CS:4470  Health Data Analytics  3
CS:4510  Human-Computer Interaction for Computer Science  3
CS:4630  Mobile Computing  3
CS:4700/MATH:4860  High Performance and Parallel Computing  3
CS:5630  Cloud Computing Technology  3
MATH:4840  Mathematics of Machine Learning  3
STAT:3210  Experimental Design and Analysis  3
STAT:4520/IGPI:4522/PSQF:4520  Bayesian Statistics  3
STAT:4560  Statistics for Risk Modeling I  3
STAT:5810/BIOS:5310/IGPI:5310  Research Data Management  3
Other advanced computer science or statistics courses approved by advisor

Capstone Courses

Course #  Title  Hours
Both of these:
DATA:4880  Data Science Creative Component  1
DATA:4890  Data Science Practicum  2

The Department of Statistics and Actuarial Science and the Department of Computer Science collaborate to offer the major in data science.

Honors

Honors in the Major

Students majoring in data science have the opportunity to graduate with honors in the major. They must maintain a grade-point average (GPA) of at least 3.67 in their major and a cumulative University of Iowa GPA of at least 3.33. Students must complete an honors thesis.

Students are responsible for finding a faculty member willing to supervise their honors project. The faculty member must approve the proposed project and a timetable for the work. Credit for thesis work must be earned in either CS:3990 Honors in Computer Science or Informatics for work supervised by a computer science faculty member or an honors course supervised by a statistics and actuarial science faculty member.

Honors in data science also satisfies the 12 s.h. experiential learning requirement for University of Iowa honors students.

University of Iowa Honors Program

In addition to honors in the major, students have opportunities for honors study and activities through membership in the University of Iowa Honors Program. Visit Honors at Iowa to learn about the university’s honors program.

Membership in the UI Honors Program is not required to earn honors in the data science major.

Career Advancement

Today, nearly every business, government, social media platform, and educational institution collects and analyzes data about its users, logistics and operations, and media presence in the hope of extracting valuable insights and utilizing the resulting efficiencies.

As an example, Amazon is the company most closely identified with a data-driven business model. Starting just over 25 years ago as an online book seller with a relatively crude crowdsourced book review platform and simple recommender system technology, it was subsequently augmented with extensive tracking of customer page views, advertising hits, data about prior purchases, and an aggressive emphasis on data-driven operational efficiencies.

Amazon has become the major player in U.S. retail and a prime example of the strategic value of big data.

Data science graduates may pursue careers as data scientists. This position allows them to apply their understanding of statistics, as well as algorithm and software design, to create and develop the next generation of data analysis tools.
The Pomerantz Career Center offers multiple resources to help students find internships and jobs.

### Academic Plans

#### Four-Year Graduation Plan

The Four-Year Graduation Plan is not available to students majoring in data science. Students work with their advisors on individual graduation plans.

#### Sample Plan of Study

Sample plans represent one way to complete a program of study. Actual course selection and sequence will vary and should be discussed with an academic advisor. For additional sample plans, see MyUI.

## Data Science, BS

### Course Title Hours

#### Academic Career

**First Year**

**Fall**

- CS:1210 Computer Science I: Fundamentals 4
- ENGL:1200 or RHET:1030 The Interpretation of Literature or Rhetoric 3 - 4
- MATH:1550 Engineering Mathematics I: Single Variable Calculus 4
- GE CLAS Core: World Languages First Level Proficiency or elective course C 4 - 5
- CSI:1600 Success at Iowa 2

**Hours** 17-19

**Spring**

- ENGL:1200 or RHET:1030 The Interpretation of Literature or Rhetoric 3 - 4
- STAT:2010 Discrete Methods and Computing 3
- CS:2210 Discrete Structures 3
- MATH:1560 Engineering Mathematics II: Multivariable Calculus 4
- GE CLAS Core: World Languages Second Level Proficiency or elective course C 4 - 5

**Hours** 17-19

#### Second Year

**Fall**

- STAT:3200 Applied Linear Regression 3
- CS:2230 Computer Science II: Data Structures 4
- GE CLAS Core: Diversity and Inclusion D 3
- GE CLAS Core: Natural Sciences without Lab D 3
- GE CLAS Core: World Languages Third Level Proficiency or elective course C 4 - 5

**Hours** 17-18

**Spring**

- CS:3330 Algorithms 3
- MATH:2700 Introduction to Linear Algebra 4
- GE CLAS Core: International and Global Issues D 3
- GE CLAS Core: World Languages Fourth Level Proficiency or elective course C 4 - 5

**Hours** 17-18

#### Third Year

**Fall**

- STAT:3100 Introduction to Mathematical Statistics I F 3
- CS:4400 Database Systems 3
- GE CLAS Core: Natural Sciences with Lab D 4
- GE CLAS Core: Social Sciences D 3
- Elective course E 3

**Hours** 16

**Spring**

- STAT:4540 or CS:5430 Statistical Learning G or Machine Learning 3
- STAT:3101 Introduction to Mathematical Statistics II H 3
- STAT:4580 Data Visualization and Data Technologies H 3
- GE CLAS Core: Literary, Visual, and Performing Arts D 3
- Elective course E 3

**Hours** 15

#### Fourth Year

**Fall**

- DATA:4880 Data Science Creative Component 1
- Major: advanced elective I course I 3
- Major: advanced elective II course I 3
- GE CLAS Core: Historical Perspectives D 3
- Elective course E 3

**Hours** 13

**Spring**

- DATA:4890 Data Science Practicum 2
- Major: advanced elective III course I 3
- GE CLAS Core: Values and Culture D 3
- Elective course E 1
- Elective course E 3
- Degree Application: apply on MyUI before deadline (typically in February for spring, September for fall) J

**Hours** 12

**Total Hours** 124-130

---

a Sustainability must be completed by choosing a course that has been approved for Sustainability AND for one of these General Education areas: Natural Sciences; Quantitative and Formal Reasoning; Social Sciences; Historical Perspectives; International and Global Issues; Literary, Visual, and Performing Arts; or Values and Culture.

b Enrollment in math courses requires completion of a placement exam.

c Students who have completed four years of a single language in high school have satisfied the GE CLAS Core World Languages requirement. Enrollment in world languages courses requires a placement exam, unless enrolling in a first-semester-level course.

d GE CLAS Core courses may be completed in any order unless used as a prerequisite for another course. Students should consult with an advisor about the best sequencing of courses.

e Students may use elective courses to earn credit towards the total s.h. required for graduation or to complete a double major, minors, or certificates.
f Typically this course is offered in fall semesters only. Check MyUI for course availability since offerings are subject to change.
g Typically STAT:4540 is offered in fall semesters only and CS:5430 is offered in spring semesters only. Check MyUI for course availability since offerings are subject to change.
h Typically this course is offered in spring semesters only. Check MyUI for course availability since offerings are subject to change.
i Students should select at least one computer science course and one statistics course for their advanced electives.
j Please see Academic Calendar, Office of the Registrar website for current degree application deadlines. Students should apply for a degree for the session in which all requirements will be met. For any questions on appropriate timing, contact your academic advisor or Graduation Services.