Data Science

Overview
Data Science is a rapidly growing academic discipline fueled by the proliferation of rich and complex data emerging from activities in science, industry, and governments. As a result, there is strong demand for data science professionals today in Iowa as well as across the nation and globe, and this market is expected to continue to grow in the next decade. The data science programs are intended for students who wish to study the data science discipline for its own sake as well as for students studying any discipline at Iowa State University with the goal of enabling them to work in data science. The courses in the data science program are designed to provide students with the requisite background that would enable them to take jobs with significant data science components, e.g., establishing and operating data analysis pipelines. The capstone will provide an opportunity for students to apply data science concepts to a domain problem while working in a multi-disciplinary team setting.

The Data Science major is intended for students with strong quantitative backgrounds and has the goal of educating students on the technical fundamentals of Data Science, with a focus on developing the knowledge and skills needed to transform data into insights. The major is an excellent opportunity for individuals who want to prepare themselves for the exciting Data Scientist positions that are in high demand today.

The minor in Data Science is intended for students studying any discipline at Iowa State and is designed to give students an in-depth understanding of data science as it is applied to a variety of domains.

The certificate in Data Science is intended for students studying any discipline at Iowa State and is designed to prepare them for future work with significant data science components. The capstone will provide an opportunity for students to apply data science concepts to a domain problem while working in a multi-disciplinary team setting.

Student Learning Outcomes for Data Science Major
After successfully completing the program, students majoring in Data Science will demonstrate

1. an understanding of and an ability to apply the following data science concepts, tools and methods to data analysis pipelines:
   a. data acquisition
   b. data preprocessing
   c. exploratory data analysis
   d. inferential and predictive thinking, modeling and analysis
   e. computational thinking, data structures, and algorithms
2. an understanding of ethical, legal, societal, and economic concerns in the application of data science concepts
3. an ability to visualize, interpret and communicate the output of data analysis pipelines to stakeholders
4. an ability to function on multi-disciplinary teams using concepts and tools from data science

See Undergraduate Minor and Undergraduate Certificate subpages for the respective learning outcomes.

Data Science Major
Purpose
The Bachelor of Science degree program in Data Science is intended for students with strong quantitative backgrounds and has the goal of educating students on the technical fundamentals of data sciences, with a focus on developing the knowledge and skills needed to manage and analyze large-scale, heterogeneous data to address a wide range of problems.

Requirements
The B.S. in Data Science consists of 120 total credit hours including:
1. 39 credit hours in the major core, three credits of which constitute a capstone course that is expected to provide experiential learning; (2) 9 credit hours in an application emphasis area to examine applications and theory of data sciences in a specific area; and (3) 23 credit hours of foundation courses. The capstone course will provide an opportunity for students to apply data science concepts to an application area while working in a multi-disciplinary team setting.

Data Science Major Requirements

<table>
<thead>
<tr>
<th>Course Code</th>
<th>Course Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS 110</td>
<td>Orientation to Data Science</td>
<td>R</td>
</tr>
<tr>
<td>DS 201</td>
<td>Introduction to Data Science</td>
<td>3</td>
</tr>
<tr>
<td>DS 202</td>
<td>Data Acquisition and Exploratory Data Analysis</td>
<td>3</td>
</tr>
<tr>
<td>DS 303</td>
<td>Concepts and Applications of Machine Learning</td>
<td>3</td>
</tr>
<tr>
<td>DS 401</td>
<td>Data Science Capstone</td>
<td>3</td>
</tr>
<tr>
<td>COM S 228</td>
<td>Introduction to Data Structures</td>
<td>3</td>
</tr>
<tr>
<td>COM S 230</td>
<td>Discrete Computational Structures</td>
<td>3</td>
</tr>
<tr>
<td>or CPR E 310</td>
<td>Theoretical Foundations of Computer Engineering</td>
<td></td>
</tr>
<tr>
<td>COM S 311</td>
<td>Introduction to the Design and Analysis of Algorithms</td>
<td>3</td>
</tr>
<tr>
<td>COM S 363</td>
<td>Introduction to Database Management Systems</td>
<td>3</td>
</tr>
<tr>
<td>CPR E 419</td>
<td>Software Tools for Large Scale Data Analysis</td>
<td>4</td>
</tr>
<tr>
<td>STAT 301</td>
<td>Intermediate Statistical Concepts and Methods</td>
<td>4</td>
</tr>
<tr>
<td>STAT 347</td>
<td>Probability and Statistical Theory for Data Science</td>
<td>4</td>
</tr>
<tr>
<td>STAT 477</td>
<td>Introduction to Categorical Data Analysis</td>
<td>3</td>
</tr>
</tbody>
</table>

At least 9 credits from the following categories to fulfill the emphasis area:
At least 6 credits from courses at the 300, 400, or 500 level from the following designations:

A B E, ADVRT, AGRON, AN S, ARCH, BCBIO, COM S, C R P, CPR E, CYB E, CYBSC, DS, E E, ECON, FIN, GIS, J E, JL MC, LA, LING, MATH, MIS, MKT, POL S, S E, SOC, STAT, TSM

At least 3 credits from one of the following courses:

BCBIO 401 Bioinformatics of Sequences
BCBIO 406 Bioinformatics of OMICS
COM S 342 Principles of Programming Languages
COM S 413 Program Analysis
COM S 421 Logic for Mathematics and Computer Science
COM S 424 Introduction to High Performance Computing
COM S 426 Introduction to Parallel Algorithms and Programming
COM S 435 Algorithms for Large Data Sets: Theory and Practice
COM S 440 Principles and Practice of Compiling
COM S 454 Distributed Systems
COM S 461 Principles and Internals of Database Systems
COM S 474 Introduction to Machine Learning
CPR E 388 Embedded Systems II: Mobile Platforms
CPR E 416 Software Evolution and Maintenance
CPR E 425 High Performance Computing for Scientific and Engineering Applications
CPR E 431 Basics of Information System Security
E E 425 Machine learning: A Signal Processing Perspective
DS 490X Independent Study
DS 490HX Independent Study
I E 312 Optimization
I E 483 Data Mining
I E 487 Big Data Analytics and Optimization
MATH 373 Introduction to Scientific Computing
MATH 407 Applied Linear Algebra
MATH 424 Introduction to High Performance Computing
MATH 481 Numerical Methods for Differential Equations
STAT 471 Introduction to Experimental Design
STAT 473 Introduction to Survey Sampling
STAT 475 Introduction to Multivariate Data Analysis

Total Credits: 9

Toward satisfying pre-requisites and requirements of the College of Liberal Arts and Sciences, the following courses or their equivalents are also required:

COM S 127 Introduction to Computer Programming 3
COM S 227 Object-oriented Programming 4
MATH 165 Calculus I 4
MATH 166 Calculus II 4
MATH 265 Calculus III 4
MATH 207 Matrices and Linear Algebra 3
STAT 201 Introduction to Statistical Concepts and Methods 4

World Language 3 years in high school or 1 year in college 0 - 8

Natural Science 8

Social Science 9

Arts and Humanities 12

LAS 203 Professional Career Preparation 1

The following courses meet the communication proficiency requirement:

LIB 160 Introduction to College Level Research 1
ENGL 150 Critical Thinking and Communication 3
ENGL 250 Written, Oral, Visual, and Electronic Composition 3

One of the following:

ENGL 302 Business Communication 3
ENGL 314 Technical Communication 3
ENGL 332 Visual Communication of Quantitative Information (cross-listed as STAT 332) 3

As majors in the College of Liberal Arts and Sciences, Data Science students must meet College of Liberal Arts and Sciences (http://catalog.iastate.edu/collegeofliberalartsandsciences/ #lascollegerequirementstext) and University-wide requirements (http://catalog.iastate.edu/collegescurricula/) for graduation in addition to those stated above for the major.

LAS majors require a minimum of 120 credits, including a minimum of 45 credits at the 300/400 level or above taken at a four-year college. All such credits, including courses taken on a pass/not pass basis, may be used to meet this requirement. You must also complete the LAS world language requirement and career proficiency requirement.

Students in all ISU majors must complete a three-credit course in U.S. diversity and a three-credit course in international perspectives. Check (http://www.registrar.iastate.edu/courses/div-ip-guide.html) for a list of approved courses. Discuss with your advisor how the two courses that you select can be applied to your graduation plan.

According to the university-wide Communication Proficiency Grade Requirement, students must demonstrate their communication proficiency by earning a grade of C or better in ENGL 250. The Data
Science program requires a C or higher in the upper-level ENGL course (302, 314, or 332).

**Four Year Plan**

**B.S., Data Science**

**Freshman**

<table>
<thead>
<tr>
<th>Fall</th>
<th>Credits</th>
<th>Spring</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS 110</td>
<td>R MATH 166</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>MATH 165</td>
<td>4 COM S 227</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>COM S 127</td>
<td>3 STAT 201</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>ENGL 150</td>
<td>3 ENGL 250</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>LIB 160</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Science</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>14</strong></td>
<td></td>
<td><strong>15</strong></td>
</tr>
</tbody>
</table>

**Sophomore**

<table>
<thead>
<tr>
<th>Fall</th>
<th>Credits</th>
<th>Spring</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS 201</td>
<td>3 DS 202</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>MATH 265</td>
<td>4 MATH 207</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>COM S 228</td>
<td>3 COM S 230 or CPR E 310</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>STAT 301</td>
<td>4 Social Science/International Perspectives</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Arts and Humanities/U.S. Diversity</td>
<td>3 Arts and Humanities</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>LAS 203</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>17</strong></td>
<td></td>
<td><strong>16</strong></td>
</tr>
</tbody>
</table>

**Junior**

<table>
<thead>
<tr>
<th>Fall</th>
<th>Credits</th>
<th>Spring</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS 303</td>
<td>3 COM S 363</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>STAT 347</td>
<td>4 STAT 477</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>COM S 311</td>
<td>3 Arts and Humanities</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Arts and Humanities (300+ level)</td>
<td>3 Natural Science</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Elective or World Language</td>
<td>3-4 Elective or World Language</td>
<td>3-4</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>16-17</strong></td>
<td></td>
<td><strong>16-17</strong></td>
</tr>
</tbody>
</table>

**Senior**

<table>
<thead>
<tr>
<th>Fall</th>
<th>Credits</th>
<th>Spring</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Emphasis Area</td>
<td>3 DS 401</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Application Emphasis Area</td>
<td>3 CPR E 419</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>ENGL 302, 314, or 332</td>
<td>3 Application Emphasis Area</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Natural Science</td>
<td>4 Social Science (300+ Level)</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>13</strong></td>
<td></td>
<td><strong>13</strong></td>
</tr>
</tbody>
</table>

The major elective courses will come from Data Science emphasis areas as outlined on the Undergraduate Major page. A student must take at least 9 credits of 300+ courses from a wide variety of designations. Additionally, 3 credits must be chosen from a list of elective courses with a data or analysis emphasis.

All students are required to take at least 45 hours of courses at the 300+ level or above. This may require taking additional electives.

**Data Science Minor**

**Purpose**

The minor in data science is intended for students studying any discipline at Iowa State and is designed to give students an in-depth understanding of data science as it is applied to a variety of domains. The minor in data science will prepare students with the technical and communication skills to enter the workforce as domain experts with data science skills.

**Learning Outcomes for Data Science Minor**

After completing the minor in data science, students will demonstrate:

- an ability to apply data science concepts, tools and technologies to data analysis pipelines,
- an understanding of ethical, legal, societal, and economic concerns in application of data science concepts,
- an ability to visualize, interpret and communicate the output of data analysis pipelines to stakeholders, and
- an ability to function on multi-disciplinary teams using concepts and tools from data science.

**Requirements**

The minor in data science requires the completion of 15 credit hours, including 9 credits from the data science core and 6 credits from approved data science electives.

At least 6 credits in courses numbered 300-level or above taken at ISU with a grade of C or higher.

At least 9 credits used for the minor cannot be used to meet any other department, college or university requirement for the baccalaureate degree except to satisfy the total credit requirement for graduation and to meet credit requirements in courses numbered 300 or above.

Courses for the minor cannot be taken on a pass/not-pass basis.

**Course Requirements for Data Science Minor**

Core Courses (9 credits)

<table>
<thead>
<tr>
<th>Course</th>
<th>Title</th>
<th>Credits</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS 201</td>
<td>Introduction to Data Science (Required)</td>
<td>3</td>
</tr>
<tr>
<td>DS 202</td>
<td>Data Acquisition and Exploratory Data Analysis (Required)</td>
<td>3</td>
</tr>
<tr>
<td>DS 301</td>
<td>Applied Data Modeling and Predictive Analysis (Required)</td>
<td>3</td>
</tr>
</tbody>
</table>
DS 301 has a prerequisite of an introductory statistics course: STAT 101, STAT 104, STAT 105, STAT 201, STAT 226, STAT 231, STAT 305, STAT 322, or STAT 330.

Electives (6 credits)

A B E 316 Applied Numerical Methods for Agricultural and Biosystems Engineering 3
ADVRT 335 Advertising Media Planning 3
ADVRT 497J Ad Tech 3
AGRON 270 Geospatial Technologies 3
AGRON 425 Crop and Soil Modeling 3
BCBIO 322 Introduction to Bioinformatics and Computational Biology 3
COM S 311 Introduction to the Design and Analysis of Algorithms 3
COM S 363 Introduction to Database Management Systems 3
COM S 424 Introduction to High Performance Computing 3
COM S 435 Algorithms for Large Data Sets: Theory and Practice 3
COM S 453 Privacy Preserving Algorithms and Data Security 3
COM S 474 Introduction to Machine Learning 3
C R P 251 Fundamentals of Geographic Information Systems 3
C R P 351 Intermediate Geographic Information Systems 3
C R P 452 Geographic Data Management and Planning Analysis 3
C R P 454 Fundamentals of Remote Sensing and Spatial Analysis 3
C R P 456 GIS Programming and Automation 3
CPR E 419 Software Tools for Large Scale Data Analysis 4
CPR E 426 Introduction to Parallel Algorithms and Programming 4
DS 490X Independent Study 4
ECON 371 Introductory Econometrics 4
E E 428X Image Analysis from Machine Learning 3
ENGL 332 Visual Communication of Quantitative Information 3
FIN 450 Analytical Methods in Finance 3
I E 312 Optimization 3
I E 483 Data Mining 3
L A 558 Web Mapping and Spatial Data Visualization 3
LING 410 Language as Data 3
MATH 304 Combinatorics 3
MATH 314 Graph Theory 3
MATH 373 Introduction to Scientific Computing 3
MATH 422X Mathematical Principles of Data Science 3
MIS 436 Introduction to Business Analytics 3
MIS 446 Advanced Business Analytics 3
MKT 368 Marketing Analytics 3
STAT 301 Intermediate Statistical Concepts and Methods 4
STAT 330 Probability and Statistics for Computer Science 3
STAT 475 Introduction to Multivariate Data Analysis 3
STAT 477 Introduction to Categorical Data Analysis 3
STAT 483 Empirical Methods for the Computational Sciences 3
STAT 486 Introduction to Statistical Computing 3
TSM 433 Precision Agriculture 3

Data Science Certificate

Purpose

The certificate in data science is intended for students studying any discipline at Iowa State and is designed to prepare them for future work with significant data science components. The data science certificate is also available to students who have already earned a Baccalaureate degree from Iowa State or elsewhere. The capstone will provide an opportunity for students to apply data science concepts to a domain problem while working in a multi-disciplinary team setting. The certificate in data science will prepare students with the technical and communication skills to enter the workforce as domain experts with data science skills.

Learning Outcomes for Data Science Certificate

After completing the certificate in data science, students will demonstrate:

- an ability to apply data science concepts, tools and technologies to data analysis pipelines,
- an understanding of ethical, legal, societal, and economic concerns in application of data science concepts,
- an ability to visualize, interpret and communicate the output of data analysis pipelines to stakeholders, and
- an ability to function on multi-disciplinary teams using concepts and tools from data science.

Requirements

The certificate in data science requires the completion of 21 credit hours, including 9 credits from the data science core, 9 credits from approved data science electives, and a three-credit data science capstone experience.

At least 9 credits must be taken in courses numbered at the 300-level or above.

At least 9 credits used for the certificate cannot be used to meet any other department, college or university requirement for the baccalaureate degree.
degree except to satisfy the total credit requirement for graduation and to meet credit requirements in courses numbered 300 or above.

Courses for the certificate cannot be taken on a pass/not-pass basis.

**Course Requirements for Data Science Certificate**

**Core Courses (9 credits)**

- DS 201  Introduction to Data Science (Required)  3
- DS 202  Data Acquisition and Exploratory Data Analysis (Required)  3
- DS 301  Applied Data Modeling and Predictive Analysis (Required)  3

* DS 301 has a prerequisite of an introductory statistics course: STAT 101, STAT 104, STAT 105, STAT 201, STAT 226, STAT 231, STAT 305, STAT 322, or STAT 330.

**Electives (9 credits)**

- A B E 316  Applied Numerical Methods for Agricultural and Biosystems Engineering  3
- ADVRT 335  Advertising Media Planning  3
- ADVRT 497J  Ad Tech  3
- BCBIO 322  Introduction to Bioinformatics and Computational Biology  3
- COM S 311  Introduction to the Design and Analysis of Algorithms  3
- COM S 363  Introduction to Database Management Systems  3
- COM S 424  Introduction to High Performance Computing  3
- COM S 435  Algorithms for Large Data Sets: Theory and Practice  3
- COM S 453  Privacy Preserving Algorithms and Data Security  3
- COM S 474  Introduction to Machine Learning  3
- C R P 251  Fundamentals of Geographic Information Systems  3
- C R P 351  Intermediate Geographic Information Systems  3
- C R P 452  Geographic Data Management and Planning Analysis  3
- C R P 456  GIS Programming and Automation  3
- CPR E 419  Software Tools for Large Scale Data Analysis  4
- CPR E 426  Introduction to Parallel Algorithms and Programming  4
- ECON 371  Introductory Econometrics  4
- ENGL 332  Visual Communication of Quantitative Information  3
- FIN 450  Analytical Methods in Finance  3
- I E 312  Optimization  3
- I E 483  Data Mining  3
- LING 410  Language as Data  3
- MATH 304  Combinatorics  3
- MATH 314  Graph Theory  3
- MATH 373  Introduction to Scientific Computing (MATH 422X::Mathematical Principals of Data Science)  3
- MATH 422X  Mathematical Principals of Data Science  3
- MIS 436  Introduction to Business Analytics (::Mathematical Principals of Data Science)  3
- MIS 446  Advanced Business Analytics  3
- MKT 368  Marketing Analytics  3
- STAT 301  Intermediate Statistical Concepts and Methods  4
- STAT 330  Probability and Statistics for Computer Science  3
- STAT 475  Introduction to Multivariate Data Analysis  3
- STAT 477  Introduction to Categorical Data Analysis  3
- STAT 483  Empirical Methods for the Computational Sciences  3
- STAT 486  Introduction to Statistical Computing  3
- DS 401  Data Science Capstone  3