# DATA SCIENCE

Data Science is a field of study that combines computer science (programming, databases, and algorithms) and statistical methodology, both with a strong mathematical foundation, to apply to diverse areas in ethical ways. Data scientists work in many areas, including business, economics, medicine, epidemiology, agriculture, environmental sciences, sports, and all aspects of government. With the increasing digitization and networking of society, data have become ever more ubiquitous, further expanding the demand for data scientists and their expertise in the collection, management, and analysis of data.

## About this Program

- **College:** Liberal Arts and Sciences (http://catalog.ufl.edu/UGRD/colleges-schools/UGLAS/)
- **Degrees:** Bachelor of Science
- **Credits for Degree:** 120
- **Contact:** Email (dathien@stat.ufl.edu)
- **More Info**

*To graduate with this major, students must complete all university, college, and major requirements.*

### Department Information

The mission of the Department of Statistics is to provide its students with a fundamental understanding of statistical reasoning and methodology, to train them to apply this knowledge to the collection and analysis of data, and to prepare them for careers in a highly technological society in which science and decision-making are increasingly driven by a rapid expansion in the quantity and availability of data.
**Website (https://stat.ufl.edu/)**

### CONTACT

Email (staff@stat.ufl.edu) | 352.392.1941 (tel) | 352.392.5175 (fax)

P.O. Box 118545
102 GRIFFIN-FLOYD HALL
GAINESVILLE FL 32611-8545
Map (http://campusmap.ufl.edu/#/index/0010)

#### Curriculum

- Actuarial Science Minor
- Combination Degrees
- Data Analytics Certificate
- Data Science
- Statistics
- Statistics Minor

Data Science majors draw inference from large data generated from a variety of disciplines. Core courses cover mathematical foundations of data science, programming, algorithms, and databases as well as statistical methods for data science. Majors will also learn about data science in practice within subject matter areas.

Students who wish to major in data science must consult a department advisor early in their programs.

## Requirements for the Major

The BS in Data Science requires a minimum of 64 credits in data science and related coursework plus 7 credits of required foundation coursework. Students must receive minimum grades of B (within 2 attempts including withdrawals) in STA2023 or STA 3032 as well as MAC 2312 and MAC 2313. All other required core and major elective courses must be completed with a minimum grade of C within two attempts (including withdrawals).

- It is important that the prerequisites of each course are met before the course is attempted.
- To take STA 3100, which is required for the major, students must meet the prerequisite: (STA 2023 with a minimum grade of B) or (STA 3032 with a minimum grade of B) or (AP Statistics with a score of 4 or higher out of 5).
- Students cannot retake core or statistics elective courses after earning a minimum grade of C, with the exception of STA 2023 or STA 3032, MAC 2312 and MAC 2313, in which students must receive a minimum grade of B.
- The grades from all attempts to satisfy core requirements will be used to compute the minimum GPA.

A minimum of 18 credits of major coursework must be taken at UF, including a minimum of 12 credits of core coursework.

## Required Foundation Coursework

| Code | Title | Credits |
| --- | --- | --- |
| MAC 2311 | Analytic Geometry and Calculus 1 | 4 |
| STA 2023 | Introduction to Statistics 1 | 3 |
| or STA 3032 | Engineering Statistics | |
| **Total Credits** | | **7** |

## Required Major Coursework

| Code | Title | Credits |
| --- | --- | --- |
| **Data Science Major Core Coursework** | | |
| **Mathematics** | | **17** |
| MAC 2312 | Analytic Geometry and Calculus 2 (minimum grade of B) | |
| MAC 2313 | Analytic Geometry and Calculus 3 (minimum grade of B) | |
| MAD 2502 | Intro to Computational Math | |
| MAS 3114 | Computational Linear Algebra | |
| MAS 4115 | Linear Algebra for Data Science | |
| **Statistics** | | **18** |
| STA 3100 | Programming With Data in R | |
| STA 4321 | Introduction to Probability | |
| STA 4322 | Introduction to Statistics Theory | |
| STA 4210 | Regression Analysis | |
| STA 4241 | Statistical Learning in R | |
| STA 4273 | Statistical Computing in R | |
| **Computer Science** | | **17** |
| COP 3502C | Programming Fundamentals 1 (MAC 2311 coreq) | |
| COP 3503C | Programming Fundamentals 2 | |
| MAD 3107 | Discrete Mathematics | |
| or COT 3100 | Applications of Discrete Structures | |
| COP 3530 | Data Structures and Algorithm | |
| CIS 4301 | Information and Database Systems 1 | |
| **Ethics** | | **3** |
| PHI 3681 | Ethics, Data, and Technology | |
| **Subject Area Electives** | | **9** |
| Select 3 | | |
| ANT 4930 | Special Topics in Anthropology | |
| CAP 3032 | Interactive Modeling and Animation 1 | |
| CGS 4144 | Introduction to Bioinformatic Algorithms | |
| CLA 3700 | Classical Archaeology | |
| ECO 4422 | Econometrics 2 | |
| EXP 4174C | Laboratory in Sensory Processes | |
| GIS 3043 | Foundations of Geographic Information Systems | |
| INR 4931 | Special Topics in International Relations | |
| LIN 4071 | Intro to Corpus Linguistics | |
| LIN 4702C | Methods in Psycholinguistics | |
| PCB 4085 | Genetical Ethics | |
| PHC 3793 | Higher Thinking for Healthy Humans: AI in Healthcare and Public Health | |
| PHZ 3152 | Advanced Computational Techniques | |
| POS 4931 | Special Topics in Political Science | |
| PSB 4343C | Laboratory in Cognitive Neuroscience | |
| REL 2104 | Environmental Ethics | |
| REL 3082 | Global Ethics | |
| REL 3160 | Religion and Science | |
| STA 4930 | Special Topics | |
| STA 4956 | Overseas Studies | |
| SYD 4020 | Population | |
| SYD 4021 | US Population Issues | |
| WST 3610 | Gender, Race and Science | |
| WST 3703 | History of American Medicine: Race, Class, Gender, and Science | |
| WST 4002 | Data Feminisms | |
| WST 4704 | Discrimination and Health | |

## Relevant Minors and Certificates

Data Science majors may want to consider a minor in actuarial science, which prepares students for careers as actuaries. Required courses cover the material for the beginning examinations and VEE credits leading to an associateship in the major national actuarial societies

### Critical Tracking

Critical Tracking records each student's progress in courses that are required for progress toward each major. Please note the critical-tracking requirements below on a per-semester basis.

For degree requirements outside of the major, refer to CLAS Degree Requirements: Structure of a CLAS Degree (http://catalog.ufl.edu/UGRD/colleges-schools/UGLAS/#degreerequirementstext).

Equivalent critical-tracking courses as determined by the State of Florida Common Course Prerequisites (https://cpm.flvc.org/advance-search/) may be used for transfer students.

## Semester 1

- Complete MAC 2311
- 2.0 UF GPA Required

## Semester 2

- Complete MAC 2312 with a minimum grade of B and (STA 2023 or STA 3032). Minimum grade of B required for STA 3100 prerequisite.
- 2.0 UF GPA Required

## Semester 3

- Complete MAC 2313 with a minimum grade of B and MAD 2502
- 2.0 UF GPA Required

## Semester 4

- Complete COP 3502C and MAS 3114  and  STA 3100
- 2.5 critical-tracking GPA required
- 2.0 UF GPA Required

## Semester 5

- Complete COP 3503C and (MAD 3107 or COT 3100) and STA 4210 and STA 4321
- 2.5 critical-tracking GPA required
- 2.0 UF GPA Required

## Semester 6

- Complete COP 3530 and MAS 4115 and PHI 3681 and STA 4322
- 2.0 UF GPA Required

## Semesters 7-8

- Complete CIS 4301 and STA 4241 and STA 4273 and one Subject Area elective
- 2.0 UF GPA Required

### Model Semester Plan

Students are expected to complete the Writing, Civic Literacy, summer enrollment, and Quest requirements while in the process of taking the courses below. Students are also expected to complete the general education international (GE-N) requirements concurrently with another general education requirement (typically, GE-C, H, or S) as part of the CLAS Basic Distribution requirements. One of the two general education mathematics courses must be a pure math course.

College of Liberal Arts and Sciences allows students additional flexibility in its Distribution Requirements. Students may count a maximum of 6 credits TOTAL from the CLAS Distribution course lists towards Humanities, Social and Behavioral Sciences, or Biological and Physical Sciences, with no more than 3 credits of Humanities, 3 credits of Social and Behavioral Sciences, or 6 credits of Biological or Physical Sciences.

The full list of major-specific requirements for this major can be found on the Overview tab. College of Liberal Arts and Sciences degree requirements can be found on the College's degree requirements page (https://catalog.ufl.edu/UGRD/colleges-schools/UGLAS/#degreerequirementstext).

MAC 2312, MAC 2313, the Subject Area electives, and the Data Ethics course count towards 3000-level or above electives outside of this multidisciplinary major.

To remain on track, students must complete the appropriate critical-tracking courses, which appear in bold. These courses must be completed by the terms as listed above in the Critical Tracking criteria.

*This semester plan represents an example progression through the major. Actual courses and course order may be different depending on the student's academic record and scheduling availability of courses. Prerequisites still apply.*

| Course | Title | Credits |
|---|---|---|
| **Semester One** | | |
| Quest 1 (Gen Ed Humanities, if needed) | | 3 |
| MAC 2311 | Analytic Geometry and Calculus 1 (**Critical Tracking**; State Core Gen Ed Mathematics) | 4 |
| State Core Gen Ed Biological or Physical Sciences (http://catalog.ufl.edu/UGRD/academic-programs/general-education/#genedcoursestext) | | 3 |
| State Core Gen Ed Composition | | 3 |
| | **Credits** | **13** |
| **Semester Two** | | |
| MAC 2312 | Analytic Geometry and Calculus 2 (**Critical Tracking**; Gen Ed Mathematics; minimum grade of B) | 4 |
| MAD 2502 | Intro to Computational Math (**Critical Tracking**) | 3 |
| STA 2023 | Introduction to Statistics 1 (**Critical Tracking**; minimum grade of B) | 3 |
| or STA 3032 | or Engineering Statistics | |
| Gen Ed Composition | | 3 |
| Gen Ed Physical Sciences | | 3 |
| | **Credits** | **16** |
| **Semester Three** | | |
| Quest 2 (Gen Ed Biological or Physical Sciences-area not taken in semester one) | | 3 |
| MAC 2313 | Analytic Geometry and Calculus 3 (**Critical Tracking**; Gen Ed Mathematics; minimum grade of B) | 4 |
| State Core Gen Ed Social Science | | 3 |
| CLAS Foreign Language Proficiency Requirement [1] | | 4-5 |
| | **Credits** | **14-15** |
| **Semester Four** | | |
| COP 3502C | Programming Fundamentals 1 (**Critical Tracking**) | 4 |
| MAS 3114 | Computational Linear Algebra (**Critical Tracking**) | 3 |
| STA 3100 | Programming With Data in R (**Critical Tracking**) | 3 |
| Gen Ed Social and Behavioral Sciences | | 3 |
| CLAS Foreign Language Proficiency Requirement [1] | | 3-5 |
| | **Credits** | **16-18** |
| **Semester Five** | | |
| COP 3503C | Programming Fundamentals 2 (**Critical Tracking**) | 4 |
| MAD 3107 | Discrete Mathematics (**Critical Tracking**) | 3 |
| or COT 3100 | or Applications of Discrete Structures | |
| STA 4210 | Regression Analysis (**Critical Tracking**) | 3 |
| STA 4321 | Introduction to Probability (**Critical Tracking**) | 3 |
| Elective (or CLAS Foreign Language Proficiency Requirement if 4-3-3 language option) [1] | | 3 |
| | **Credits** | **16** |
| **Semester Six** | | |
| COP 3530 | Data Structures and Algorithm (**Critical Tracking**) | 3 |
| MAS 4115 | Linear Algebra for Data Science (**Critical Tracking**) | 3 |
| PHI 3681 | Ethics, Data, and Technology (**Critical Tracking**) | 3 |
| STA 4322 | Introduction to Statistics Theory (**Critical Tracking**; ) | 3 |
| Gen Ed Humanities | | 3 |
| | **Credits** | **15** |
| **Semester Seven** | | |
| CIS 4301 | Information and Database Systems 1 (**Critical Tracking**) | 3 |
| STA 4241 | Statistical Learning in R (**Critical Tracking**) | 3 |
| Subject Area Elective (**Critical Tracking**; 3000-level or higher | | 3 |
| State Core Gen Ed Humanities (http://catalog.ufl.edu/UGRD/academic-programs/general-education/#genedcoursestext) | | 3 |
| Gen Ed Biological Sciences | | 3 |

| | | |
|---|---|---|
| Natural Science Laboratory [1] | | 1 |
| | **Credits** | **16** |
| **Semester Eight** | | |
| STA 4273 | Statistical Computing in R (**Critical Tracking**) | 3 |
| Subject Area Electives (**Critical Tracking**; 3000-level or higher) | | 6 |
| Gen Ed Social and Behavioral Sciences | | 3 |
| Elective | | 2 |
| | **Credits** | **14** |
| | **Total Credits** | **120** |

[1]     https://catalog.ufl.edu/UGRD/colleges-schools/UGLAS/#degreerequirementstext

The Data Science major enables students to achieve proficiency in the fundamentals of programming, databases, and statistical reasoning. Through coursework and projects, students will gain knowledge in problem solving, data science applications and ethics, and statistical inference. Emphasis is on developing the ability to approach real world problems and through the use of computing and statistical methods to draw valid scientific inferences.

# Before Graduating Students Must

- Complete an exam on the fundamentals of data science, which will be 5% of their grade in STA 4241 .
- Complete a data analysis project, which will be 10% of their grade in STA 4241 .
- Complete requirements for the baccalaureate degree, as determined by faculty.

# Students in the Major Will Learn to

## Student Learning Outcomes | SLOs

### Content
1. Identify, define, and describe concepts and issues in data science, including those involved in computing and programming, databases, ethics, mathematical foundations, and statistical methods.

### Critical Thinking
2. Identify sources of variability and bias in a given set of data and formulate and carefully program an appropriate statistical analysis.

### Communication
3. Clearly and effectively present ideas in speech and in writing concerning issues in the proper analysis of data.

## Curriculum Map

*I = Introduced; R = Reinforced; A = Assessed*

| Courses | SLO 1 | SLO 2 | SLO 3 |
|---|---|---|---|
| MAS 3114 | I | | |
| MAS 4115 | R | | |
| STA 3100 | I | I | I |
| STA 4321 | I | | |
| STA 4322 | I | | |
| STA 4210 | R | R | R |
| STA 4241 | A | A | A |
| STA 4273 | R | R | R |
| COP 3502C | I | | |
| COP 3503C | R | | |
| COP 3530 | I | | |
| CIS 4301 | R | | |
| PHI 3681 | I | R | R |

# Assessment Types

- Exams
- Projects
- Written and oral presentations