STATISTICAL DATA SCIENCE (BS)

The Statistical Data Science major gives students a broad training in the following core areas of data science:

- computer programming and data management
- basic and advanced data analysis
- data visualization
- data ethics

Students with this major obtain a Bachelor of Science (B.S.) degree. The major can be tailored for a student's interest in a domain concentration.

Admission

Students with this major obtain a Bachelor of Science (B.S.) degree. The major can be tailored for a student's interest in a domain concentration.

In order to apply to the Statistical Data Science major, students must have:

- A GPA of 3.2 or higher in the following classes: MATH 1132Q Calculus II, STAT 1000Q Introduction to Statistics I/STAT 1100Q Elementary Concepts of Statistics, and an introductory programming course (CSE 1010 Introduction to Computing for Engineers, CSE 1729 Introduction to Principles of Programming, or STAT 2255 Statistical Programming).
- completed at least 24 credits, 15 of which must be at the University of Connecticut, with a cumulative GPA of 3.2 or higher.

After entry into the majors, students must maintain a 3.2 cumulative GPA.

Location

- Storrs Campus

Modality

- Online

Requirements

~~Students receiving a B.S. in Statistical Data Science are required to take 36 major credits.~~

Core Area Requirements

Course	Title	Credits
STAT 3255	Introduction to Data Science	3
Programming and Data Management		
STAT 2255	Statistical Programming	3
or ECON 3322	Open Source Programming with Python for Economists	
Basic Data Analysis		
STAT 3215Q	Applied Linear Regression in Data Science	3
Select one of the following:		3
STAT 3025Q	Statistical Methods	
STAT 3375Q	Introduction to Mathematical Statistics I	
MATH 3160	Probability	
Data Ethics		
Select one of the following:		3
PHIL 3202	Data Ethics	
STAT 3205	Data Management, Programming, and Privacy	
Data Visualization		
Select at least three credits of the following:		3-4
STAT 3675Q	Statistical Computing [1]	
EEB 4100	Big Data Science for Biologists [2]	
GSCU 3510	Cartographic Techniques	
Advanced Analysis		
MATH 2210Q	Applied Linear Algebra	3
STAT 4255	Introduction to Statistical Learning	3
Capstone Course		
STAT 4915	Data Science in Action [3]	3
Writing in the Major		
STAT 4916W	Writing in Data Science [3]	1
Domain Concentration Sequence		
Students must complete one of the nine-credit domain concentration sequences listed below		9
Total Credits		**37-38**

1. Students completing a Statistics domain concentration must take STAT 3375Q Introduction to Mathematical Statistics I and STAT 3675Q Statistical Computing to meet these requirements.
2. Recommended for students completing the Biological Data Science domain concentration.
3. Students completing a Biological Data Science domain concentration may take any of the following to meet the capstone and W requirement:
   1. STAT 4915 Data Science in Action, STAT 4916W Writing in Data Science,
   2. EEB 4896W Senior Research Thesis in Ecology and Evolutionary Biology, or
   3. MCB 4897W Senior Research Thesis or MCB 4997W Senior Honors Research Thesis.

   Credits in EEB 4896W Senior Research Thesis in Ecology and Evolutionary Biology cannot simultaneously count towards both an Honors thesis in EEB and a Data Science capstone. Separate thesis requirements are needed in MCB 4997W (https://catalog.uconn.edu/search/?P=EEB%204896W) Senior Honors Research Thesis for students pursuing Honors Scholar or Laureate designations in both Molecular and Cell Biology and Statistical Data Science.

Domain Concentration Sequence

To complete the domain concentration sequence, students must take at least nine credits from one of the following groups:

Advanced Statistics

Course	Title	Credits
STAT 3445	Introduction to Mathematical Statistics II	3
Select two of the following:		6
STAT 3515Q	Design of Experiments	
STAT 4625	Introduction to Biostatistics	
STAT 4825	Applied Time Series	
STAT 4845	Applied Spatio-Temporal Statistics	
STAT 4865	Deep Learning	

STAT 4190	Field Study Internship [1]
STAT 4195	Special Topics [2]

1 At least and no more than three credits of STAT 4190 Field Study Internship may count towards the major and must be pre-approved by the Department of Statistics for adequate data science content.

2 STAT 4195 Special Topics may count at most once towards the domain with the consent of the advisor or undergraduate program director dependent on topic.

American Political Representation

Course	Title	Credits
Select three of the following:		9
POLS 2607	American Political Parties	
POLS 2618	Politics of Inequality	
POLS 3608	The Art, Science, and Business of Political Campaigns	
POLS 3612	Electoral Behavior	
POLS 3617	American Political Economy	
POLS 3625	Public Opinion	

Biological Data Science

Course	Title	Credits
Select three of the following: [3]		9-12
EEB 3899	Independent Study [1]	
MCB 3201	Gene Expression	
MCB 3421	Introduction to Molecular Evolution and Bioinformatics	
MCB 3637	Practical Methods in Microbial Genomics	
MCB 3895	Special Topics [2]	
MCB 4008	Techniques of Biophysical Chemistry	
MCB 4009	Structure and Function of Biological Macromolecules	
MCB 4014	Structure and Dynamics of Macromolecular Complexes	
MCB 4896	Undergraduate Research [1]	
MCB 4996	Honors Undergraduate Research [1]	

1 Only three credits of EEB 3899 Independent Study, MCB 4896 Undergraduate Research, or MCB 4996 Honors Undergraduate Research can count towards the major, and these credits cannot simultaneously count towards another major or degree.

2 MCB 3895 Special Topics may count at most once towards the domain with the consent of the Biological Data Science domain advisor dependent on topic.

3 The graduate courses listed below may also count towards the Biological Data Science domain: EEB 5050 Fundamentals of Ecological Modeling, EEB 5300 Practical Genomics in Ecology and Evolution, EEB 5348 Population Genetics, EEB 5349 Phylogenetics, MCB 5430 Analysis of Eukaryotic Functional Genomic Data, MCB 5631 Sequence-based Microbial Community Analysis. Prerequisites vary; please consult instructor.

Students can choose any three courses[1] from the list above based on availability, however, interested students might consider choosing subsets of courses from the list above that align with established sub-areas:

Course	Title	Credits
Genome Sequencing and Analysis		
MCB 3201	Gene Expression	3
MCB 3421	Introduction to Molecular Evolution and Bioinformatics	4
MCB 3637	Practical Methods in Microbial Genomics	3
Molecular Structure and Function		
MCB 4008	Techniques of Biophysical Chemistry	3
MCB 4009	Structure and Function of Biological Macromolecules	3
MCB 4014	Structure and Dynamics of Macromolecular Complexes	3

1 Only three credits of EEB 3899 Independent Study, MCB 4896 Undergraduate Research, or MCB 4996 Honors Undergraduate Research can count towards the major, and these credits cannot simultaneously count towards another major or degree.

Financial Analysis

Course	Title	Credits
Select three of the following:		9
ECON 3313	Elementary Economic Forecasting	
ECON 3315	Financial Econometrics	
ECON 3422	International Finance	
ECON 4323	Convex Optimization with Python	

Marine Science

Course	Title	Credits
Select three of the following:		9-11
MARN 2801WE	Marine Sciences and Society	
MARN 3001	Foundations of Marine Sciences	
MARN 3002	Foundations of Marine Sciences	
MARN 3014	Marine Biology	
MARN 4001	Measurement and Analysis in Coastal Ecosystems	
MARN 4010	Biological Oceanography	
MARN 4210Q	Experimental Design in Marine Ecology	

Population Dynamics

Course	Title	Credits
Select three of the following:		9
SOCI 2110W	Sociology of Education	
SOCI 2651/2651W	Sociology of Families	
SOCI 2660/2660W	Sociology of Health	
SOCI 2820	Sociological Perspectives on Poverty	
SOCI 2901/2901W	Urban Sociology	
SOCI 3751	Population in a Changing World	

Bachelor of Science Requirements

For a Statistical Data Science major that leads to a Bachelor of Science degree, students must take:

Course	Title	Credits
Statistics Course		
Select one of the following:		4
STAT 1000Q	Introduction to Statistics I	
STAT 1100Q	Elementary Concepts of Statistics	
MATH Courses		
Selection one of the following MATH sequences:		8
Sequence 1		
MATH 1131Q & MATH 1132Q	Calculus I and Calculus II	
Sequence 2		
MATH 2141Q & MATH 2142Q	Advanced Calculus I and Advanced Calculus II	
Select one of the following:		3-4
MATH 2110Q	Multivariable Calculus	
MATH 2210Q	Applied Linear Algebra	
MATH 2410Q	Elementary Differential Equations	
Science Sequence		
Select a sequence in one of the following that includes laboratory measurements [1]		8-10
Biology (https://catalog.uconn.edu/undergraduate/liberal-arts-sciences/statistical-data-science-bs/#biology)		
BIOL 1107	Principles of Biology I	
BIOL 1108	Principles of Biology II	
or BIOL 1110	Introduction to Botany	
Chemistry		
Select one of the following sequences:		
CHEM 1124Q & CHEM 1125Q & CHEM 1126Q	Fundamentals of General Chemistry I and Fundamentals of General Chemistry II and Fundamentals of General Chemistry III	
CHEM 1127Q & CHEM 1128Q	General Chemistry I and General Chemistry II	
CHEM 1147Q & CHEM 1148Q	Honors General Chemistry I and Honors General Chemistry II	
Physics		
Select one of the following sequences:		
PHYS 1201Q & PHYS 1202Q	General Physics I and General Physics II	
PHYS 1401Q & PHYS 1402Q	General Physics with Calculus I and General Physics with Calculus II	
PHYS 1501Q & PHYS 1502Q	Physics for Engineers I and Physics for Engineers II	
PHYS 1601Q & PHYS 1602Q	Fundamentals of Physics I and Fundamentals of Physics II	
Total Credits		**23-26**

[1] For students following the former General Education curriculum, one of these courses may be used to fulfill the CA 3 laboratory science requirement of the University's General Education requirements. In addition, students must take one other CA 3 course from a different subject area, but it need not be a lab course. For students following the Common Curriculum, one of these courses may be used to fulfill the TOI 6 laboratory science requirement of the University's Common Curriculum requirements. In addition, students must take one other TOI 6 course from a different subject area, but it need not be a lab course.

Learning Objectives

1. Demonstrate proficiency in computer programming and data management, including topics such as data types, control flow, object-oriented programming, algorithmic thinking, efficient implementation of different data structures, control and data abstraction.
2. Demonstrate proficiency in expositing for technical and non-technical audiences in writing, including organizing and describing data and analyses at a level appropriate for one's audience, thinking critically about the results and implications of such analyses, creating informative and effective visualizations, and communicating an analysis pipeline through well-written, reproducible code and report.
3. Demonstrate proficiency in understanding data ethics, including systematic approaches to assessing ethical issues; privacy and confidentiality; defining research and the responsibilities associated with conducting ethical research; and implicit and structural biases in data collection and analysis.
4. Demonstrate proficiency in the domain concentration.