

Data Science

Bachelor of Arts (BA)

The Data Science Major degree program combines computational and inferential reasoning to draw conclusions based on data about some aspect of the real world. Data scientists come from all walks of life, all areas of study, and all backgrounds. They share an appreciation for the practical use of mathematical and scientific thinking and the power of computing to understand and solve problems for business, research, and societal impact.

The Data Science Major will equip students to draw sound conclusions from data in context, using knowledge of statistical inference, computational processes, data management strategies, domain knowledge, and theory. Students will learn to carry out analyses of data through the full cycle of the investigative process in scientific and practical contexts. Students will gain an understanding of the human and ethical implications of data analytics and integrate that knowledge in designing and carrying out their work.

The Data Science major requirements include "Data 8" (STAT C8) and "Data 100" (COMPSCI C100), the core lower-division and upper-division elements of the major, along with courses from each of the following requirement groups:

- Foundations in Mathematics and Computing
- Computational and Inferential Depth
- Modeling, Learning and Decision Making
- Probability
- Human Contexts and Ethics
- Domain Emphasis

All students will select a Domain Emphasis, a cluster of one lower division course and two upper division courses, that brings them into the context of a domain and allows them to build bridges with data science.

Declaring the Major

Students can apply to declare the Data Science major after completing all the lower-division prerequisites (see the Major Requirements tab). For applicants with prerequisites in progress, applications will be reviewed after the grades for all prerequisites are available.

It is necessary for applicants to achieve a minimum prerequisite grade point average (GPA) in order to declare the Data Science major. Information on this GPA and the process to apply for admission to the major can be found on the Declaring the Major (<http://data.berkeley.edu/degrees/data-science-ba/declaring>) web page.

Minor Program

We expect that a minor in Data Science will be available to students from all Colleges, except for students with a Data Science major. The proposal for the minor program is currently in development. Check the Data Science program website (<https://data.berkeley.edu>) for details and status updates.

In addition to the University, campus, and college requirements, listed on the College Requirements tab, students must fulfill the below requirements specific to the major program. As the program is new, these

requirements may undergo changes. Please check the Data Science program website (<http://data.berkeley.edu/degrees>) for updates.

General Guidelines

- All courses taken to fulfill the major requirements below must be taken for graded credit, other than courses listed which are offered on a Pass/No Pass basis only. Other exceptions to this requirement are noted as applicable.
- Only one upper-division course may be used to simultaneously fulfill requirements for a student's major and minor programs. No more than two upper-division courses can overlap between two majors.
- A minimum grade point average (GPA) of 2.0 must be maintained in both upper and lower division courses used to fulfill the major requirements.

Lower Division Prerequisites

STAT/COMPSCI/ Foundations of Data Science ("Data 8") ¹ INFO C8		4
MATH 1A & MATH 1B	Calculus and Calculus	8
MATH 54	Linear Algebra and Differential Equations	4
or STAT 89A	Linear Algebra for Data Science	
or EECS 16A & EECS 16B	Designing Information Devices and Systems I and Designing Information Devices and Systems II	
or PHYSICS 89	Introduction to Mathematical Physics	
COMPSCI 61A	The Structure and Interpretation of Computer Programs	4
or COMPSCI 8	Computational Structures in Data Science	
or ENGIN 7	Introduction to Computer Programming for Scientists and Engineers	
COMPSCI 61B	Data Structures	4

¹ For a transitional period through Spring 2019, students may substitute STAT C8 Foundations of Data Science ("Data 8") with STAT 20, STAT 21, or STAT 135 when they have previously taken COMPSCI 61A. More details about this "grandfathering" policy available on our website: <http://data.berkeley.edu/degrees/data-science-ba>. (<http://data.berkeley.edu/degrees/data-science-ba>)

In some cases, students may complete alternative courses to satisfy the above prerequisites. See the lower-division requirements page on the Data Science program website (<http://data.berkeley.edu/degrees>) for more details.

Lower Division Requirements

Students will also be required to take one lower division course towards their choice of Domain Emphasis. See the Domain Emphasis tab for more information.

Upper Division Requirements

Students will be required to complete eight unique upper-division courses for a total of 28 or more units from the following six requirement categories. Please see each category below for the specific course options available and the number of courses required to fulfill that category.

Principles and techniques of data science

Students will be required to complete "Data 100", Principles & Techniques of Data Science (COMPSCI C100) for 4 units.

COMPSCI/STAT	Principles & Techniques of Data Science C100	4
--------------	----------------------------------------------	---

Computational and Inferential Depth

Students will be required to take two upper division courses comprising 7 or more units that provide computational and inferential depth beyond that provided in Data 100 (COMPSCI C100) and their lower-division courses.

Choose two courses comprising 7+ units from the following:

ASTRON 128	Astronomy Data Science Laboratory	4
COMPSCI 161	Computer Security	4
COMPSCI 162	Operating Systems and System Programming	4
COMPSCI 164	Programming Languages and Compilers	4
COMPSCI 168	Introduction to the Internet: Architecture and Protocols	4
COMPSCI 169	Software Engineering	4
COMPSCI 170	Efficient Algorithms and Intractable Problems	4
COMPSCI 186	Introduction to Database Systems	4
COMPSCI 188	Introduction to Artificial Intelligence	4
ECON 140	Economic Statistics and Econometrics	4
or ECON 141	Econometric Analysis	
EECS 127	Optimization Models in Engineering	4
EL ENG 120	Signals and Systems	4
EL ENG 123	Digital Signal Processing	4
EL ENG 129	Course Not Available	3
ESPM 174	Design and Analysis of Ecological Research	4
IND ENG 115	Industrial and Commercial Data Systems	3
IND ENG 135	Applied Data Science with Venture Applications	3
IND ENG 173	Introduction to Stochastic Processes	3
INFO 154	Data Mining and Analytics	3
INFO 159	Natural Language Processing	4
INFO 190	Special Topics in Information (Introduction to Data Visualization)	1-3
NUC ENG 175	Methods of Risk Analysis	3
PHYSICS 188	Bayesian Data Analysis and Machine Learning for Physical Sciences (previously PHYSICS 188)	4
STAT 135	Concepts of Statistics	4
STAT 150	Stochastic Processes	3
STAT 151A	Linear Modelling: Theory and Applications	4
STAT 152	Sampling Surveys	4
STAT 153	Introduction to Time Series	4
STAT 158	The Design and Analysis of Experiments	4
STAT 159	Reproducible and Collaborative Statistical Data Science	4

Probability

Students will be required to take one upper-division course on probability.

Choose one of the following:

EL ENG 126	Probability and Random Processes	4
------------	----------------------------------	---

IND ENG 172	Probability and Risk Analysis for Engineers	3
STAT 134	Concepts of Probability	4
STAT 140	Probability for Data Science	4

Modeling, Learning, and Decision-Making

Students will be required to take one upper-division course on modeling, learning, and decision-making.

Choose one of the following:

COMPSCI 182	Designing, Visualizing and Understanding Deep Neural Networks	4
COMPSCI 189	Introduction to Machine Learning	4
IND ENG 142	Introduction to Machine Learning and Data Analytics	3
STAT 102	Course Not Available ("Data 102")	4
STAT 154	Modern Statistical Prediction and Machine Learning	4

Human Contexts and Ethics

Students will be required to take one course from a curated list of courses that establish a human, social, and ethical context in which data analytics and computational inference play a central role.

AFRICAM 134	Information Technology and Society	4
or AFRICAM/AMERSTD C134	Information Technology and Society	
BIO ENG 100	Ethics in Science and Engineering	3
CY PLAN 101	Introduction to Urban Data Analytics	4
HISTORY C184D/STS C104D	Human Contexts and Ethics of Data - DATA/History/STS	4
INFO 188	Behind the Data: Humans and Values	3
ISF 100J	The Social Life of Computing	4
PHILOS 121	Moral Questions of Data Science	4
PB HLTH C160/ESPM C167	Environmental Health and Development	4

Domain Emphasis

Students will also be required to take two upper division course towards their choice of Domain Emphasis. See the Domain Emphasis tab for more information.

Domain Emphases that students can choose from (list up-to-date as of April 2019):

- Applied Mathematics, Modeling, and Optimization (<https://data.berkeley.edu/degrees/domain-emphasis/applied-math-and-modeling>)
- Business and Industrial Analytics (<https://data.berkeley.edu/degrees/domain-emphasis/business-and-industrial-analytics>)
- Cognition (<https://data.berkeley.edu/degrees/domain-emphasis/cognition>)
- Computational Biology Methods (<https://data.berkeley.edu/degrees/domain-emphasis/computational-biology-methods>)
- Data Arts and Humanities (<https://data.berkeley.edu/degrees/domain-emphasis/data-arts-and-humanities>)

- Ecology and the Environment (<https://data.berkeley.edu/degrees/domain-emphasis/ecology-and-environment>)
- Economics (<https://data.berkeley.edu/degrees/domain-emphasis/economics>)
- Environment, Resource Management, and Society (<https://data.berkeley.edu/degrees/domain-emphasis/environment-resources-society>)
- Evolution and Biodiversity (<https://data.berkeley.edu/degrees/domain-emphasis/evolution-and-biodiversity>)
- Geospatial Information and Technology (<https://data.berkeley.edu/degrees/domain-emphasis/geospatial-info-and-technology>)
- Human Biology (<https://data.berkeley.edu/degrees/domain-emphasis/human-biology>)
- Human and Population Health (<https://data.berkeley.edu/degrees/domain-emphasis/human-and-population-health>)
- Human Behavior and Psychology (<https://data.berkeley.edu/degrees/domain-emphasis/human-behavior-and-psych>)
- Inequalities in Society (<https://data.berkeley.edu/degrees/domain-emphasis/inequalities-in-society>)
- Linguistic Sciences (<https://data.berkeley.edu/degrees/domain-emphasis/linguistic-sciences>)
- Molecular Biology and Genomics (<https://data.berkeley.edu/degrees/domain-emphasis/molecular-bio-and-genomics>)
- Neurosciences (<https://data.berkeley.edu/degrees/domain-emphasis/neurosciences>)
- Organizations and the Economy (<https://data.berkeley.edu/degrees/domain-emphasis/organizations-and-the-economy>)
- Physical Science Analytics (<https://data.berkeley.edu/degrees/domain-emphasis/physical-science-analytics>)
- Quantitative Social Science (<https://data.berkeley.edu/degrees/domain-emphasis/quantitative-social-science>)
- Robotics (<https://data.berkeley.edu/degrees/domain-emphasis/robotics>)
- Social Welfare, Health, and Poverty (<https://data.berkeley.edu/degrees/domain-emphasis/social-welfare-health-poverty>)
- Social Policy and Law (<https://data.berkeley.edu/degrees/domain-emphasis/social-policy-and-law>)
- Sustainable Development and Engineering (<https://data.berkeley.edu/degrees/domain-emphasis/sustainable-dev-and-engineering>)
- Urban Science (<https://data.berkeley.edu/degrees/domain-emphasis/urban-science>)

Domain Emphases give students a grounded understanding of a particular domain of data-intensive research, relevant theory, or an integrative intellectual thread. A Domain Emphasis is comprised of three courses chosen from a list. Each Domain Emphasis is rooted in a lower division course, which is typically also a prerequisite for the upper division courses.

A Domain Emphasis is not limited to courses that are intended to be specifically for data science. Rather, they should bring the data science student into the context of a domain. That may involve understanding the vocabulary, methods of study, theoretical foundations, or cultural outlook of the domain. The student needs to become able to build the bridges with data science in carrying out the emphasis, rather than expecting each course to do it for them.

Students will select one course from a short list of lower-division prerequisites, and two courses from a list of upper-division courses.

The lower division course is a required element of the Domain Emphasis.

What to think about when selecting a domain emphasis:

- Courses you take for the 7-course L&S Breadth requirement may fulfill the lower-division course for a domain. Even if you don't yet know which domain to choose, your breadths may work for you for the major as well.
- Allow yourself some flexibility—choose three or four upper-division courses from the course list (<https://data.berkeley.edu/degrees/domain-emphasis>) that you'd like to take, rather than two, so you can make sure you continue making progress if you are unable to take a particular course.
- Be advised that it is important to examine whether seats are typically available in the courses for the domain you select—many courses in other departments have priority enrollment groups.

Some domain emphases are currently undergoing revision and are subject to change. Course lists for each of the domain emphases are available on the Data Science program website. (<https://data.berkeley.edu/degrees/domain-emphasis>)

List of Domain Emphases

(List accurate as of April 2019):

- Applied Mathematics, Modeling, and Optimization (<https://data.berkeley.edu/degrees/domain-emphasis/applied-math-and-modeling>)
- Business and Industrial Analytics (<https://data.berkeley.edu/degrees/domain-emphasis/business-and-industrial-analytics>)
- Cognition (<https://data.berkeley.edu/degrees/domain-emphasis/cognition>)
- Computational Biology Methods (<https://data.berkeley.edu/degrees/domain-emphasis/computational-biology-methods>)
- Data Arts and Humanities (<https://data.berkeley.edu/degrees/domain-emphasis/data-arts-and-humanities>)
- Ecology and the Environment (<https://data.berkeley.edu/degrees/domain-emphasis/ecology-and-environment>)
- Economics (<https://data.berkeley.edu/degrees/domain-emphasis/economics>)
- Environment, Resource Management, and Society (<https://data.berkeley.edu/degrees/domain-emphasis/environment-resources-society>)
- Evolution and Biodiversity (<https://data.berkeley.edu/degrees/domain-emphasis/evolution-and-biodiversity>)
- Geospatial Information and Technology (<https://data.berkeley.edu/degrees/domain-emphasis/geospatial-info-and-technology>)
- Human Biology (<https://data.berkeley.edu/degrees/domain-emphasis/human-biology>)
- Human and Population Health (<https://data.berkeley.edu/degrees/domain-emphasis/human-and-population-health>)
- Human Behavior and Psychology (<https://data.berkeley.edu/degrees/domain-emphasis/human-behavior-and-psych>)
- Inequalities in Society (<https://data.berkeley.edu/degrees/domain-emphasis/inequalities-in-society>)

- Linguistic Sciences (<https://data.berkeley.edu/degrees/domain-emphasis/linguistic-sciences>)
- Molecular Biology and Genomics (<https://data.berkeley.edu/degrees/domain-emphasis/molecular-bio-and-genomics>)
- Neurosciences (<https://data.berkeley.edu/degrees/domain-emphasis/neurosciences>)
- Organizations and the Economy (<https://data.berkeley.edu/degrees/domain-emphasis/organizations-and-the-economy>)
- Physical Science Analytics (<https://data.berkeley.edu/degrees/domain-emphasis/physical-science-analytics>)
- Quantitative Social Science (<https://data.berkeley.edu/degrees/domain-emphasis/quantitative-social-science>)
- Robotics (<https://data.berkeley.edu/degrees/domain-emphasis/robotics>)
- Social Welfare, Health, and Poverty (<https://data.berkeley.edu/degrees/domain-emphasis/social-welfare-health-poverty>)
- Social Policy and Law (<https://data.berkeley.edu/degrees/domain-emphasis/social-policy-and-law>)
- Sustainable Development and Engineering (<https://data.berkeley.edu/degrees/domain-emphasis/sustainable-dev-and-engineering>)
- Urban Science (<https://data.berkeley.edu/degrees/domain-emphasis/urban-science>)

Undergraduate students must fulfill the following requirements in addition to those required by their major program.

For detailed lists of courses that fulfill college requirements, please review the College of Letters & Sciences (<http://guide.berkeley.edu/archive/2019-20/undergraduate/colleges-schools/letters-science>) page in this Guide. For College advising appointments, please visit the L&S Advising (<https://ls.berkeley.edu/advising/about-undergraduate-advising-services>) Pages.

University of California Requirements

Entry Level Writing (<http://writing.berkeley.edu/node/78>)

All students who will enter the University of California as freshmen must demonstrate their command of the English language by fulfilling the Entry Level Writing requirement. Fulfillment of this requirement is also a prerequisite to enrollment in all reading and composition courses at UC Berkeley.

American History and American Institutions (<http://guide.berkeley.edu/archive/2019-20/undergraduate/colleges-schools/letters-science/american-history-institutions-requirement>)

The American History and Institutions requirements are based on the principle that a US resident graduated from an American university, should have an understanding of the history and governmental institutions of the United States.

Berkeley Campus Requirement

American Cultures (<http://americancultures.berkeley.edu/students/courses>)

All undergraduate students at Cal need to take and pass this course in order to graduate. The requirement offers an exciting intellectual environment centered on the study of race, ethnicity and culture of the United States. AC courses offer students opportunities to be part of research-led, highly accomplished teaching environments, grappling with the complexity of American Culture.

College of Letters & Science Essential Skills Requirements

Quantitative Reasoning (<http://guide.berkeley.edu/archive/2019-20/undergraduate/colleges-schools/letters-science/quantitative-reasoning-requirement>)

The Quantitative Reasoning requirement is designed to ensure that students graduate with basic understanding and competency in math, statistics, or computer science. The requirement may be satisfied by exam or by taking an approved course.

Foreign Language (<http://guide.berkeley.edu/archive/2019-20/undergraduate/colleges-schools/letters-science/foreign-language-requirement>)

The Foreign Language requirement may be satisfied by demonstrating proficiency in reading comprehension, writing, and conversation in a foreign language equivalent to the second semester college level, either by passing an exam or by completing approved course work.

Reading and Composition (<http://guide.berkeley.edu/archive/2019-20/undergraduate/colleges-schools/letters-science/reading-composition-requirement>)

In order to provide a solid foundation in reading, writing, and critical thinking the College requires two semesters of lower division work in composition in sequence. Students must complete parts A & B reading and composition courses by the end of their second semester and a second-level course by the end of their fourth semester.

College of Letters & Science 7 Course Breadth Requirements

Breadth Requirements (<http://guide.berkeley.edu/archive/2019-20/undergraduate/colleges-schools/letters-science/#breadthrequirementstext>)

The undergraduate breadth requirements provide Berkeley students with a rich and varied educational experience outside of their major program. As the foundation of a liberal arts education, breadth courses give students a view into the intellectual life of the University while introducing them to a multitude of perspectives and approaches to research and scholarship. Engaging students in new disciplines and with peers from other majors, the breadth experience strengthens interdisciplinary connections and context that prepares Berkeley graduates to understand and solve the complex issues of their day.

Unit Requirements

- 120 total units
- Of the 120 units, 36 must be upper division units
- Of the 36 upper division units, 6 must be taken in courses offered outside your major department

Residence Requirements

For units to be considered in "residence," you must be registered in courses on the Berkeley campus as a student in the College of Letters & Science. Most students automatically fulfill the residence requirement by attending classes here for four years. In general, there is no need to be concerned about this requirement, unless you go abroad for a semester or year or want to take courses at another institution or through UC Extension during your senior year. In these cases, you should make

an appointment to meet an adviser to determine how you can meet the Senior Residence Requirement.

Note: Courses taken through UC Extension do not count toward residence.

Senior Residence Requirement

After you become a senior (with 90 semester units earned toward your BA degree), you must complete at least 24 of the remaining 30 units in residence in at least two semesters. To count as residence, a semester must consist of at least 6 passed units. Intercampus Visitor, EAP, and UC Berkeley-Washington Program (UCDC) units are excluded.

You may use a Berkeley Summer Session to satisfy one semester of the Senior Residence requirement, provided that you successfully complete 6 units of course work in the Summer Session and that you have been enrolled previously in the college.

Modified Senior Residence Requirement

Participants in the UC Education Abroad Program (EAP), Berkeley Summer Abroad, or the UC Berkeley Washington Program (UCDC) may meet a Modified Senior Residence requirement by completing 24 (excluding EAP) of their final 60 semester units in residence. At least 12 of these 24 units must be completed after you have completed 90 units.

Upper Division Residence Requirement

You must complete in residence a minimum of 18 units of upper division courses (excluding UCEAP units), 12 of which must satisfy the requirements for your major.

Sample plans for completing major coursework are included below. These are not comprehensive plans which will reflect the situation of every student. These sample plans are meant only to serve as a baseline guide for structuring a plan of study, and only include the minimum courses for meeting L&S college requirements, 7-course breadth, and the Data Science major requirements.

For new freshmen (four-year plan):

Freshman				
	Fall	Units	Spring	Units
STAT C8		4	COMPSCI 61A or 88	4
MATH 1A		4	MATH 1B	4
ENGLISH R1A (R&C A)		4	LINGUIS R1B (R&C B)	4
		12		12
Sophomore				
	Fall	Units	Spring	Units
COMPSCI 61B		4	MATH 54	4
PSYCH 1 (Lower-division DE)		3	PSYCH 167A (DE #1 / AC / breadth)	3
MCELLBI 32 (breadth)		3	HISTART 100 (breadth)	4
		10		11
Junior				
	Fall	Units	Spring	Units
COMPSCI C100		4	STAT 140 (Probability)	4
INFO 154 (C&ID #1)		3	PSYCH 124 (DE #2)	3
POLECON 101 (breadth)		4	MUSIC 139 (breadth)	4
		11		11

	Fall	Units	Spring	Units
COMPSCI 186 (C&ID #2)		4	STAT 102 (MLDM)	4
EPS 102 (breadth)		4	HISTORY C1E (HCE / breadth)	4
		8		4
Total Units: 79				

For transfer students (two-year plan):

**Note, this sample plan assumes students will complete calculus and two or three breadth courses at their previous college or university, which may not reflect the reality for every transfer student.*

First Year				
	Fall	Units	Spring	Units
STAT C8		4	COMPSCI 61B	4
COMPSCI 88 or 61A		2	COMPSCI C1A	4
MATH 54 or STAT 89A		4	PHILOS 121 (HCE / breadth)	4
SOCIOL 3AC (Lower-division DE / AC / breadth)		4		
		14		12
Second Year				
	Fall	Units	Spring	Units
IND ENG 172 (Probability)		3	STAT 102 (MLDM)	4
COMPSCI 170 (C&ID #1)		4	NUC ENG 17E (C&ID #2)	3
LEGALST 102 (DE #1)		4	LEGALST 123 (DE #2)	4
ISF 100A (breadth)		4	FILM 129 (breadth)	4
		15		15
Total Units: 56				

Student Teams

Each semester, we recruit dozens of students to participate in our student teams as interns and volunteers. Teams include Communications (<https://data.berkeley.edu/news/communications-team>), Analytics, External Relations, and Curriculum Development (<https://data.berkeley.edu/education/modules>). Interested students can email ds-teams@berkeley.edu with questions about the opportunities.

Data Scholars

The Data Scholars program addresses issues of underrepresentation in the data science community by establishing a welcoming, educational, and empowering environment for underrepresented and nontraditional students. The program, which offers specialized tutoring, advising, mentorship, and workshops, is especially suited for students who can bring diverse perspectives to the field of Data Science. Learn more here (<http://bit.ly/data-scholars>).

Data Peer Consulting

Students in our consulting network help make data science accessible across the broader campus community by providing technical support and tutoring. Peer consultants are available at Moffitt Library on a drop-in basis. Learn more here (<https://data.berkeley.edu/education/data-peer-consulting>).

Data Science Peer Advising

Academic Peer Advisors are available to help fellow students choose classes, explore academic interests, and learn how to declare the Data Science major. The Peer Advising services are available on a drop-in basis at Moffitt Library. Learn more here (<https://data.berkeley.edu/degrees/peer-advising>).

Discovery Research Program

The Data Science Discovery Research program connects undergraduates with hands-on, team-based opportunities to contribute to cutting-edge research projects with graduate and post-doctoral students, community impact groups, entrepreneurial ventures, and educational initiatives across UC Berkeley. Learn more here (<https://data.berkeley.edu/research/discovery>).

Data Science Nexus

The Data Science Nexus is an alliance of data science student organizations on campus that work together to build community, host industry events, and provide academic support for students. In recognition of the extraordinarily diverse and multi-faceted nature of data science, members of the Nexus come from a variety of domains. Learn more here (<https://data.berkeley.edu/data-science-student-community>).