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 C8 and DATA 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

The Minor in Data Science at UC Berkeley aims to provide students with practical knowledge of the methods and techniques of data analysis, as well as the ability to think critically about the construction and implications of data analysis and models. The minor will empower students across the wide array of campus disciplines with a working knowledge of statistics, probability, and computation that allow students not just to participate in data science projects, but to design and carry out rigorous computational and inferential analysis for their field of interest. Check the Data Science Minor program website (https://data.berkeley.edu/academics/undergraduate-programs/data-science-minor/) for details.

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. Please check the Data Science program website (https://data.berkeley.edu/academics/undergraduate-programs/data-science-major/) for updates.

General Guidelines

- All courses taken to fulfill the major requirements below must be taken for letter-graded credit.
- 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 all courses toward the major, and in all upper-division courses toward the major.

Lower Division Prerequisites

DATA C8	Foundations of Data Science	4
MATH 1A	Calculus	3-4
or MATH 10A	Methods of Mathematics: Calculus, Statistics, and Combinatorics	
or MATH 16A	Analytic Geometry and Calculus	
MATH 1B	Calculus	4
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 88	Course Not Available	
or ENGIN 7	Introduction to Computer Programming for Scientists and Engineers	
COMPSCI 61B	Data Structures	4

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 (https://data.berkeley.edu/academics/undergraduate-programs/data-science-major/) for more details.

Lower Division Requirements

Students will also be required to take one lower division course towards their choice of Domain Emphasis.

Upper Division Requirements

Students will be required to complete 8 unique upper-division courses for a total of 28 or more units from the following requirement categories.

Principles and techniques of data science

DATA C100	Principles & Techniques of Data Science	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 and the 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
or COMPSCI 169A	Introduction to Software Engineering	
or COMPSCI W169A	Software Engineering	
COMPSCI 170	Efficient Algorithms and Intractable Problems	4
COMPSCI 186	Introduction to Database Systems	4
or COMPSCI W186	Introduction to Database Systems	
COMPSCI 188	Introduction to Artificial Intelligence	4
DATA 144	Data Mining and Analytics	3
ECON 140	Econometrics	4
or ECON 141	Econometrics (Math Intensive)	
EECS 127	Optimization Models in Engineering	4
EL ENG 120	Signals and Systems	4
EL ENG 123	Digital Signal Processing	4
ENVECON C118	Introductory Applied Econometrics	4
ESPM 174	Design and Analysis of Ecological Research	4
IAS C118	Introductory Applied Econometrics	4
IND ENG 115	Industrial and Commercial Data Systems	3
IND ENG 135	Applied Data Science with Venture Applications	3
IND ENG 165	Engineering Statistics, Quality Control, and Forecasting	4
IND ENG 173	Introduction to Stochastic Processes	3
INFO 159	Natural Language Processing	4
INFO 190	Special Topics in Information (Introduction to Data Visualization - only when offered on this topic)	4
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
UGBA 147	Special Topics in Operations and Information Technology Management (Advanced Business Analytics - only when offered on this topic)	1-4

Probability

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

Choose one of the following:

DATA C140	Probability for Data Science	4
EL ENG 126	Probability and Random Processes	4
IND ENG 172	Probability and Risk Analysis for Engineers	4
STAT 134	Concepts of Probability	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
or COMPSCI L	Course Not Available	
or COMPSCI W	Course Not Available	
COMPSCI 189	Introduction to Machine Learning	4
DATA C102	Data, Inference, and Decisions	4
IND ENG 142	Introduction to Machine Learning and Data Analytics	3
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
DATA C104	Human Contexts and Ethics of Data - DATA/History/STS	4
DIGHUM 100	Theory and Method in the Digital Humanities	3
INFO 188	Behind the Data: Humans and Values	3
ISF 100J	The Social Life of Computing	4
NWMEDIA 151AC	Transforming Tech: Issues and Interventions in STEM and Silicon Valley	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 courses towards their choice of Domain Emphasis.

Domain Emphases that students can choose from:

- 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/)
- Philosophical Foundations: Evidence and Inference (https://data.berkeley.edu/academics/data-science-undergraduate-studies/data-science-major/requirements-domain-emphases/)
- Philosophical Foundations: Minds, Morals, and Machines (https://data.berkeley.edu/academics/data-science-undergraduate-studies/data-science-major/requirements-domain-emphases-0/)
- 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/)

The Minor in Data Science at UC Berkeley aims to provide students with practical knowledge of the methods and techniques of data analysis, as well as the ability to think critically about the construction and implications of data analysis and models. The minor will empower students across the wide array of campus disciplines with a working knowledge of statistics, probability, and computation that allow students not just to participate in data science projects, but to design and carry out rigorous computational and inferential analysis for their field of interest.

General Guidelines

1. All minors must be declared no later than one semester before a student's Expected Graduation Term (EGT). If the semester before EGT is fall or spring, the deadline is the last day of RRR week. If the semester before EGT is summer, the deadline is the final Friday of Summer Sessions. For more information about declaring the minor, view the Data Science minor webpage (https://data.berkeley.edu/academics/undergraduate-programs/data-science-minor/).
2. All courses for the minor must be taken for a letter grade.
3. Students must earn a C- or better in each course, and have a minimum 2.0 GPA in all courses towards the minor.
4. Students may overlap up to 1 course in the upper division requirements for the Data Science minor with each of their majors (for example, a Computer Science major may count COMPSCI/DATA/STAT C100 toward both their major and the Data Science minor).
5. A maximum of one course offered by or cross-listed with the student's major department(s) may count toward the data science minor upper-division requirements, including any overlapping course (for example, if a Computer Science major takes COMPSCI/DATA/STAT C100 toward the Data Science minor, this is the only COMPSCI, ELENG, or EECS course which may count toward the upper-division requirements for the minor).
6. An upper-division course used to fulfill a lower-division requirement (for example, Stat 134 to fulfill the probability requirement) will not be counted toward the maximum 1 course allowed to overlap with the major, nor will it fulfill one of the four upper division course requirements.
7. There is no restriction on overlap with another minor.
8. Courses used to fulfill the minor requirements may be applied toward the Seven-Course Breadth requirement, for Letters & Science students.
9. All minor requirements must be completed prior to the last day of finals during the semester in which you plan to graduate.

Lower-division Requirements

DATA C8	Foundations of Data Science	4
COMPSCI 88	Course Not Available	3-4
or COMPSCI 61A	The Structure and Interpretation of Computer Programs	
or ENGIN 7	Introduction to Computer Programming for Scientists and Engineers	
Choose one of the following: [1]		
STAT 88	Course Not Available	3-4
or COMPSCI 70	Discrete Mathematics and Probability Theory	
or MATH 10B	Methods of Mathematics: Calculus, Statistics, and Combinatorics	
or MATH 55	Discrete Mathematics	

[1] STAT 134, STAT C140, IND ENG 172, or EECS 126 may be substituted.

Upper-division Requirements

Complete a total of 4 upper-division courses in one of the following pathways:

1-Core course Pathway

DATA C100	Principles & Techniques of Data Science	4
Choose one of the following:		
AMERSTD/ AFRICAM C134	Information Technology and Society [4]	
or AFRICAM 134	Information Technology and Society	
BIO ENG 100	Ethics in Science and Engineering [3]	
CY PLAN 101	Introduction to Urban Data Analytics [4]	
DATA C104/ HISTORY C184/ STS C104D	Human Contexts and Ethics of Data - DATA/ History/STS [4]	
DIGHUM 100	Theory and Method in the Digital Humanities [3]	
ESPM C167/ PB HLTH C160	Environmental Health and Development [4]	
INFO 188	Behind the Data: Humans and Values [3]	
ISF 100J	The Social Life of Computing [4]	
NWMEDIA 151AC	Transforming Tech: Issues and Interventions in STEM and Silicon Valley [4]	
PHILOS 121	Moral Questions of Data Science [4]	

If completing the 1-core course pathway, choose TWO from the Approved Elective List **(**https://docs.google.com/document/d/1OvQls_dcVndFVqcgufywEnUzXHovrWKLLFTXa8a3VyI/edit/?usp=sharing**).**

2-core course PATHWAY

DATA C131A	Statistical Methods for Data Science	4
STAT 133	Concepts in Computing with Data	3
Choose one of the following:		
AMERSTD/ AFRICAM C134	Information Technology and Society [4]	
or AFRICAM 134	Information Technology and Society	
BIO ENG 100	Ethics in Science and Engineering [3]	
CY PLAN 101	Introduction to Urban Data Analytics [4]	
DATA C104/ HISTORY C184D/ STS C104D	Human Contexts and Ethics of Data - DATA/ History/STS [4]	
DIGHUM 100	Theory and Method in the Digital Humanities [3]	
ESPM C167/ PB HLTH C160	Environmental Health and Development [4]	
INFO 188	Behind the Data: Humans and Values [3]	
ISF 100J	The Social Life of Computing [4]	
NWMEDIA 151	Transforming Tech: Issues and Interventions in STEM and Silicon Valley [4]	
PHILOS 121	Moral Questions of Data Science [4]	

If completing the 2-core course pathway, choose ONE from the Approved Elective List **(**https://docs.google.com/document/d/1OvQls_dcVndFVqcgufywEnUzXHovrWKLLFTXa8a3VyI/edit/?usp=sharing**).**

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/undergraduate/colleges-schools/letters-science/) page in this Guide. For College advising appointments, please visit the L&S Advising (https://lsadvising.berkeley.edu/home/) 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/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/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/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/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/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 the L&S Data Science major requirements.

For new freshmen (four-year plan):

				Freshman
	Fall Units		**Spring Units**	
DATA C8	4		COMPSCI 61A or 88	3-4
MATH 1A (10A or 16A acceptable)	4		MATH 1B	4
Reading & Composition A	4		Reading & Composition B	4
DATA 88 (optional)	2		Non-major Elective	1-2
	14			**12-14**
				Sophomore
	Fall Units		**Spring Units**	
COMPSCI 61B	4		MATH 54 or STAT 89A	4
Breadth/Elective	3-4		Lower-division Domain Emphasis	3-4
Breadth/Elective	3-4		Breadth/ Elective	3-4
	10-12			**10-12**
				Junior
	Fall Units		**Spring Units**	
DATA C140 (or other approved Probability)	4		DATA C100	4
Domain Emphasis Upper-division #1			Computational & Inferential Depth #1	3-4
Breadth/Elective	3-4		Breadth/ Elective	3-4
	7-8			**10-12**
				Senior
	Fall Units		**Spring Units**	
Computational & Inferential Depth #2	3-4		DATA C102 (or other approved MLDM)	4
Domain Emphasis Upper-division #2	3-4		DATA C104 (or other approved HCE)	4
Breadth/Elective	3-4		Breadth/ Elective	3-4
	9-12			**11-12**

Total Units: 83-96

For transfer students (two-year plan):

**Note: this sample plan is based on a transfer student who has completed 1 year of calculus, linear algebra and data structures, as well as IGETC/L&S 7-Course Breadth at their previous college or university, which may not reflect the reality for every transfer student. Students should consult with a Data Science Advisor to make an individualized plan based on their specific situation.*

				First Year
	Fall Units		**Spring Units**	
DATA C8	4		DATA C100	4
COMPSCI 88 or COMPSCI 61A	3-4		DATA C140 (or other approved Probability)	4

Lower-division Domain Emphasis	2-4	American Cultures/ Upper-division Outside Major	3-4
Non-major Elective	1-2	Non-major Elective	1-2
	10-14		**12-14**
			Second Year
Fall Units		**Spring Units**	
Computational & Inferential Depth #1	3-4	DATA C102 (or other approved MLDM)	4
Domain Emphasis Upper-division #1	3-4	DATA C104 (or other approved HCE)	4
Domain Emphasis Upper-division #2	3-4	Computational & Inferential Depth #2	3-4
Non-major Elective	1-2	Non-major Elective	1-2
	10-14		**12-14**

Total Units: 44-56

Major Maps help undergraduate students discover academic, co-curricular, and discovery opportunities at UC Berkeley based on intended major or field of interest. Developed by the Division of Undergraduate Education in collaboration with academic departments, these experience maps will help you:

- **Explore** your major and gain a better understanding of your field of study
- **Connect** with people and programs that inspire and sustain your creativity, drive, curiosity and success
- **Discover** opportunities for independent inquiry, enterprise, and creative expression
- **Engage** locally and globally to broaden your perspectives and change the world
- **Reflect** on your academic career and prepare for life after Berkeley

Use the major map below as a guide to planning your undergraduate journey and designing your own unique Berkeley experience.

View the Data Science Major Map PDF. **(**https://vcue.berkeley.edu/sites/default/files/data_science.pdf**)**

Student Teams

Each semester, we recruit dozens of students to participate in our student teams as interns and volunteers, with opportunities to advance into team lead roles and other leadership positions. Teams include Communications, Operations, External Relations, and Curriculum Development (https://data.berkeley.edu/education/modules/). Interested students can email ds-teams@berkeley.edu with questions about the opportunities. Learn more here (https://data.berkeley.edu/academics/resources/student-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 office hours, consultations via appointments, and data literacy workshops open to all. Peer consultants are available at Moffitt Library on a drop-in basis or via Zoom. The team also recruits new consultants every semester. 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. Contact the Data Science Peer Advisors at ds-peer-advising@berkeley.edu. 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-program-home/).

Discovery Exchange

The Discovery Exchange (https://discoveryexchange.org/books/project-exchange/) is an online hub for Berkeley students doing independent data science projects. The Discovery Exchange will provide a platform for Berkeley students to connect with their peers to real world projects, technical training, mentoring and consulting. Learn more here (https://data.berkeley.edu/discovery-exchange/).

Data Science Nexus

Nexus highlights the interdisciplinary and intersectional nature of data science by bringing undergraduate students together through mentoring opportunities, professional development workshops, and community-building social events. Learn more here (https://data.berkeley.edu/data-science-student-community/).

Career Accelerator

The Data Science Career Accelerator program connects undergraduates with paid internship opportunities with industry-leading companies in a broad range of fields, and supports students with skill development before, during, and after the internship period. For more info contact ds-internships@berkeley.edu. Learn more here (https://data.berkeley.edu/academics/data-science-undergraduate-studies/data-science-career-accelerator/).

Expand all course descriptions [+]Collapse all course descriptions [-]

DATA 4AC Data and Justice 4 Units

Terms offered: Spring 2022, Spring 2021
This course engages students with fundamental questions of justice in relation to data and computing in American society. Data collection, visualization, and analysis have been entangled in the struggle for racial and social justice because they can make injustice visible, imaginable, and thus actionable. Data has also been used to oppress minoritized communities and institutionalize, rationalize, and naturalize systems of racial violence. The course examines key sites of justice involving data (such as citizenship, policing, prisons, environment, and health). Along with critical social science tools, students gain introductory experience and do collaborative and creative projects with data science using real-world data.
Data and Justice: Read More [+]
Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture and 1.5 hours of discussion per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Alternative to final exam.

Data and Justice: Read Less [-]

DATA 6 Introduction to Computational Thinking with Data 3 Units

Terms offered: Not yet offered
An introduction to computational thinking and quantitative reasoning, preparing students for further coursework (especially Foundations of Data Science, Data 8). Emphasizes the use of computation to gain insight about quantitative problems. Uses data from various domains in the social sciences in order to develop an understanding of the societal implications of data science. Expressions, data types, collections, and tables in Python. Programming practices, abstraction, and iteration. The data science lifecycle. Visualizing univariate and bivariate data with bar charts, histograms, plots, and maps. Introduction to statistical concepts including averages and distributions, prediction, causality, probability, sampling, and inference.
Introduction to Computational Thinking with Data: Read More [+]
Objectives & Outcomes

Course Objectives: C6 also includes quantitative reasoning concepts that aren't covered in Data 8. These include certain topics in: principles of data visualization; simulation of random processes; and understanding numerical functions through their graphs. This will help prepare students for computational and quantitative courses other than Data 8.
C6 takes advantage of the complementarity of computing and quantitative reasoning to enliven abstract ideas and build students' confidence in their ability to solve real problems with quantitative tools. Students learn computer science concepts and immediately apply them to plot functions, visualize data, and simulate random events.

Foundations of Data Science (CS/Info/Stat C8, a.k.a. Data 8) is an increasingly popular class for entering students at Berkeley. Data 8 builds students' computing skills in the first month of the semester, and students rely on these skills as the course progresses. For some students, particularly those with little prior exposure to computing, developing these skills benefits from further time and practice. C6 is a rapid introduction to Python programming, visualization, and data analysis, which will prepare students for success in Data 8.

Student Learning Outcomes: Appreciate the interdisciplinary nature of data science.
Create and use visualizations to understand univariate data and to identify associations or causal relationships in bivariate data.
Formulate questions about data and perform exploratory data analysis.
Perform basic computations in Python, and be able to work with tabular data.
Run and understand basic probabilistic simulations.
Understand the syntactic structure of Python code.
Use good practices in Python programming.

Rules & Requirements

Credit Restrictions: Students will receive no credit for DATA 6 after completing DATA C8.

Hours & Format

Summer: 6 weeks - 4 hours of lecture, 2 hours of discussion, and 4 hours of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Instructors: Yan, Hug, Harding

Formerly known as: Data Science, Undergraduate C6/Computer Science C6/Statistics C6

DATA C8 Foundations of Data Science 4 Units

Terms offered: Fall 2022, Summer 2022 8 Week Session, Spring 2022, Fall 2021, Summer 2021 8 Week Session, Fall 2020
Foundations of data science from three perspectives: inferential thinking, computational thinking, and real-world relevance. Given data arising from some real-world phenomenon, how does one analyze that data so as to understand that phenomenon? The course teaches critical concepts and skills in computer programming and statistical inference, in conjunction with hands-on analysis of real-world datasets, including economic data, document collections, geographical data, and social networks. It delves into social and legal issues surrounding data analysis, including issues of privacy and data ownership.
Foundations of Data Science: Read More [+]
Rules & Requirements

Prerequisites: This course may be taken on its own, but students are encouraged to take it concurrently with a data science connector course (numbered 88 in a range of departments)

Credit Restrictions: Students will receive no credit for DATA C8\COMPSCI C8\INFO C8\STAT C8 after completing COMPSCI 8, or DATA 8. A deficient grade in DATA C8\COMPSCI C8\INFO C8\STAT C8 may be removed by taking COMPSCI 8, COMPSCI 8, or DATA 8.

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture and 2 hours of laboratory per week

Summer: 8 weeks - 6 hours of lecture and 4 hours of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: Computer Science C8/Statistics C8/Information C8

Also listed as: COMPSCI C8/INFO C8/STAT C8

Foundations of Data Science: Read Less [-]

DATA 88 Data Science Connector 2 - 4 Units

Terms offered: Spring 2022, Spring 2021, Fall 2020
Designed to be taken in conjunction with the Foundations of Data Science (COMPSCI/INFO/STAT C8) course, each connector course will flesh out data science ideas in the context of one particular field. Blending inferential thinking and computational thinking, the course relies on the increasing availability of datasets across a wide range of human endeavor, and students' natural interest in such data, to teach students to work actively with data in a field of their interest and to interpret and critique their analyses of data. Topics vary by field, and several topics will be offered each term.
Data Science Connector: Read More [+]
Objectives & Outcomes

Course Objectives: Discuss how to formulate and substantiate an argument with evidence
Explain a variety of analytic and visualization techniques
Explore approaches to effective communication
Explore the challenges with working with primary and secondary data

Student Learning Outcomes: Apply data analysis to evaluate everyday problems
Communicate effectively in written, spoken, and graphical form about specific issues
Interpret statistical results
Know how to locate and use primary data sources
Obtain and/or collect relevant data using specific qualitative and/or quantitative research methods
Understand how to use empirical evidence to evaluate an argument

Rules & Requirements

Prerequisites: Instructors may require students to enroll concurrently or have completed Data 8 (COMPSCI/STAT/INFO C8)

Repeat rules: Course may be repeated for credit without restriction. Students may enroll in multiple sections of this course within the same semester.

Hours & Format

Fall and/or spring: 15 weeks - 2-4 hours of seminar per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Alternative to final exam.

Data Science Connector: Read Less [-]

DATA 88E Economic Models 2 Units

Terms offered: Fall 2022, Spring 2022, Fall 2021
This class aims to motivate and illustrate key concepts in economics through a series of exercises and examples that use Python Jupyter notebooks. The class covers concepts from introductory economics, microeconomic theory, econometrics, development economics, environmental economics and public economics. The course provides data science students a pathway to apply Python programming and data science concepts within the discipline of economics. The course will also gives economics students a pathway to apply programming to reinforce fundamental concepts and to advance the level of study in upper division coursework and possible thesis work.
Economic Models: Read More [+]

Objectives & Outcomes

Course Objectives: Demonstrate how to construct understanding of concepts in economics by developing and coding examples
Illustrate topics in economics through coding applications
Motivate basics of econometrics from a data science perspective

Student Learning Outcomes: Programmatically create and interpret graphs of simple equations used in microeconomics
Reason about and solve simple equations used in microeconomics through coding
Understand basic concepts in economics

Rules & Requirements

Prerequisites: You must have taken Data C8 or be concurrently enrolled in Data C8 to take this course. That being said, we are able to make exceptions if you have prior programming or data science experience; please email the course staff if you have any questions. Prior economics knowledge may be helpful but is not necessary

Hours & Format

Fall and/or spring: 15 weeks - 2 hours of lecture per week

Summer: 6 weeks - 5 hours of lecture per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Alternative to final exam.

Economic Models: Read Less [-]

DATA C88C Computational Structures in Data Science 3 Units

Terms offered: Fall 2022
Development of Computer Science topics appearing in Foundations of Data Science (C8); expands computational concepts and techniques of abstraction. Understanding the structures that underlie the programs, algorithms, and languages used in data science and elsewhere. Mastery of a particular programming language while studying general techniques for managing program complexity, e.g., functional, object-oriented, and declarative programming. Provides practical experience with composing larger systems through several significant programming projects.
Computational Structures in Data Science: Read More [+]

Objectives & Outcomes

Course Objectives: Develop a foundation of computer science concepts that arise in the context of data analytics, including algorithm, representation, interpretation, abstraction, sequencing, conditional, function, iteration, recursion, types, objects, and testing, and develop proficiency in the application of these concepts in the context of a modern programming language at a scale of whole programs on par with a traditional CS introduction course.

Student Learning Outcomes: Students will be able to demonstrate a working knowledge of these concepts and a proficiency of programming based upon them sufficient to construct substantial stand-alone programs.

Rules & Requirements

Prerequisites: MATH 1A. Also, this course is a Data Science connector course and may only be taken concurrently with or after COMPSCI C8/ DATA C8/INFO C8/STAT C8. Students may take more than one Data Science connector (88) course if they wish, concurrent with or after having taken the C8 course

Credit Restrictions: Students will receive no credit for DATA C88C after completing COMPSCI 61A.

Hours & Format

Fall and/or spring: 15 weeks - 2 hours of lecture and 2 hours of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: Computer Science 88

Also listed as: COMPSCI C88C

Computational Structures in Data Science: Read Less [-]

DATA C88S Probability and Mathematical Statistics in Data Science 3 Units

Terms offered: Fall 2022
In this connector course we will state precisely and prove results discovered while exploring data in Data 8. Topics include: probability, conditioning, and independence; random variables; distributions and joint distributions; expectation, variance, tail bounds; Central Limit Theorem; symmetries in random permutations; prior and posterior distributions; probabilistic models; bias-variance tradeoff; testing hypotheses; correlation and the regression model.
Probability and Mathematical Statistics in Data Science: Read More [+]
Rules & Requirements

Prerequisites: Prerequisite: one semester of calculus at the level of Math 16A, Math 10A, or Math 1A. Corequisite or Prerequisite: Foundations of Data Science (COMPSCI C8 / DATA C8 / INFO C8 / STAT C8)

Credit Restrictions: Students will receive no credit for STAT 88 after completing STAT 134, STAT C140, STAT 135, or STAT 102.

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture and 2 hours of discussion per week

Summer: 8 weeks - 6 hours of lecture and 4 hours of discussion per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: Statistics 88

Also listed as: STAT C88S

Probability and Mathematical Statistics in Data Science: Read Less [-]

DATA 94 Special Topics in Data Science 1 - 4 Units

Terms offered: Spring 2021
Topics will vary semester to semester.
Special Topics in Data Science: Read More [+]
Hours & Format

Fall and/or spring: 15 weeks - 1-3 hours of lecture, 0-2 hours of discussion, and 0-2 hours of laboratory per week

Summer: 8 weeks - 2-6 hours of lecture, 0-4 hours of discussion, and 0-4 hours of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required, with common exam group.

Special Topics in Data Science: Read Less [-]

DATA C100 Principles & Techniques of Data Science 4 Units

Terms offered: Fall 2022, Summer 2022 8 Week Session, Spring 2022, Fall 2021, Summer 2021 8 Week Session, Fall 2020
In this course, students will explore the data science lifecycle, including question formulation, data collection and cleaning, exploratory data analysis and visualization, statistical inference and prediction, and decision-making. This class will focus on quantitative critical thinking and key principles and techniques needed to carry out this cycle. These include languages for transforming, querying and analyzing data; algorithms for machine learning methods including regression, classification and clustering; principles behind creating informative data visualizations; statistical concepts of measurement error and prediction; and techniques for scalable data processing.
Principles & Techniques of Data Science: Read More [+]
Rules & Requirements

Prerequisites: COMPSCI C8 / DATA C8 / INFO C8 / STAT C8; and COMPSCI 61A, COMPSCI 88, or ENGIN 7; Corequisite: MATH 54 or EECS 16A

Credit Restrictions: Students will receive no credit for DATA C100\STAT C100\COMPSCI C100 after completing DATA 100. A deficient grade in DATA C100\STAT C100\COMPSCI C100 may be removed by taking DATA 100.

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture, 1 hour of discussion, and 1 hour of laboratory per week

Summer: 8 weeks - 6 hours of lecture, 2 hours of discussion, and 2 hours of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: Statistics C100/Computer Science C100

Also listed as: COMPSCI C100/STAT C100

Principles & Techniques of Data Science: Read Less [-]

DATA 101 Data Engineering 4 Units

Terms offered: Fall 2022
This course will cover the principles and practices of managing data at scale, with a focus on use cases in data analysis and machine learning. We will cover the entire life cycle of data management and science, ranging from data preparation to exploration, visualization and analysis, to machine learning and collaboration, with a focus on ensuring reliable, scalable operationalization.
Data Engineering: Read More [+]

Rules & Requirements

Prerequisites: COMPSCI 61B or INFO 206B or equivalent courses in programming with a C- or better, or Pass; AND COMPSCI C100/ DATA C100/STAT C100 or COMPSCI 189 or INFO 251 or DATA 144 or equivalent upper-division course in data science with a C- or better, or Pass

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture and 1 hour of discussion per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Instructors: Hellerstein, Jain, Parameswaran

Data Engineering: Read Less [-]

DATA C102 Data, Inference, and Decisions 4 Units

Terms offered: Fall 2022, Spring 2022, Fall 2021
This course develops the probabilistic foundations of inference in data science, and builds a comprehensive view of the modeling and decision-making life cycle in data science including its human, social, and ethical implications. Topics include: frequentist and Bayesian decision-making, permutation testing, false discovery rate, probabilistic interpretations of models, Bayesian hierarchical models, basics of experimental design, confidence intervals, causal inference, Thompson sampling, optimal control, Q-learning, differential privacy, clustering algorithms, recommendation systems and an introduction to machine learning tools including decision trees, neural networks and ensemble methods.
Data, Inference, and Decisions: Read More [+]

Rules & Requirements

Prerequisites: Mathematics 54 or Mathematics 110 or Statistics 89A or Physics 89 or both of Electrical Engineering and Computer Science 16A and Electrical Engineering and Computer Science 16B; Statistics/Computer Science C100; and any of Electrical Engineering and Computer Science 126, Statistics 140, Statistics 134, Industrial Engineering and Operations Research 172. Statistics 140 or Electrical Engineering and Computer Science 126 are preferred

Credit Restrictions: Students will receive no credit for DATA C102 after completing STAT 102, or DATA 102. A deficient grade in DATA C102 may be removed by taking STAT 102, STAT 102, or DATA 102.

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture, 1 hour of discussion, and 1 hour of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: Statistics 102

Also listed as: STAT C102

Data, Inference, and Decisions: Read Less [-]

DATA C104 Human Contexts and Ethics of Data - DATA/History/STS 4 Units

Terms offered: Fall 2022, Spring 2022, Fall 2021, Fall 2020, Spring 2020
This course teaches you to use the tools of applied historical thinking and Science, Technology, and Society (STS) to recognize, analyze, and shape the human contexts and ethics of data. It addresses key topics such as doing ethical data science amid shifting definitions of human subjects, consent, and privacy; the changing relationship between data, democracy, and law; the role of data analytics in how corporations and governments provide public goods such as health and security to citizens; sensors, machine learning and artificial intelligence and changing landscapes of labor, industry, and city life. It prepares you to engage as a knowledgeable and responsible citizen and professional in the varied arenas of our datafied world.
Human Contexts and Ethics of Data - DATA/History/STS: Read More [+]
Rules & Requirements

Credit Restrictions: Students will receive no credit for DATA C104\HISTORY C184D\STS C104D after completing DATA 104. A deficient grade in DATA C104\HISTORY C184D\STS C104D may be removed by taking DATA 104.

Hours & Format

Fall and/or spring: 15 weeks - 3-3 hours of lecture and 0-1.5 hours of discussion per week

Summer:
6 weeks - 7.5-7.5 hours of lecture and 0-3.5 hours of discussion per week
8 weeks - 6-6 hours of lecture and 0-3 hours of discussion per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: History C184D/Science and Technology Studies C104D

Also listed as: HISTORY C184D/STS C104D

Human Contexts and Ethics of Data - DATA/History/STS: Read Less [-]

DATA C131A Statistical Methods for Data Science 4 Units

Terms offered: Fall 2022, Spring 2022, Fall 2021
This course teaches a broad range of statistical methods that are used to solve data problems. Topics include group comparisons and ANOVA, standard parametric statistical models, multivariate data visualization, multiple linear regression, logistic regression and classification, regression trees and random forests. An important focus of the course is on statistical computing and reproducible statistical analysis. The course and lab include hands-on experience in analyzing real world data from the social, life, and physical sciences. The R statistical language is used.
Statistical Methods for Data Science: Read More [+]
Rules & Requirements

Prerequisites: Statistics/Computer Science/Information C8 or Statistics 20; and Mathematics 1A, Mathematics 16A, or Mathematics 10A/10B. Strongly recommended corequisite: Statistics 33A or Statistics 133

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture and 2 hours of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: Statistics 131A

Also listed as: STAT C131A

Statistical Methods for Data Science: Read Less [-]

DATA C140 Probability for Data Science 4 Units

Terms offered: Fall 2022, Spring 2022, Fall 2021, Spring 2021
An introduction to probability, emphasizing the combined use of mathematics and programming to solve problems. Random variables, discrete and continuous families of distributions. Bounds and approximations. Dependence, conditioning, Bayes methods. Convergence, Markov chains. Least squares prediction. Random permutations, symmetry, order statistics. Use of numerical computation, graphics, simulation, and computer algebra.
Probability for Data Science: Read More [+]

Objectives & Outcomes

Course Objectives: The emphasis on simulation and the bootstrap in Data 8 gives students a concrete sense of randomness and sampling variability. Stat 140 will capitalize on this, abstraction and computation complementing each other throughout.

The syllabus has been designed to maintain a mathematical level at least equal to that in Stat 134. So Stat 140 will start faster than Stat 134 (due to the Data 8 prerequisite), avoid approximations that are unnecessary when SciPy is at hand, and replace some of the routine calculus by symbolic math done in SymPy. This will create time for a unit on the convergence and reversibility of Markov Chains as well as added focus on conditioning and Bayes methods.
With about a thousand students a year taking Foundations of Data Science (Stat/CS/Info C8, a.k.a. Data 8), there is considerable demand for follow-on courses that build on the skills acquired in that class. Stat 140 is a probability course for Data 8 graduates who have also had a year of calculus and wish to go deeper into data science.

Student Learning Outcomes: Understand the difference between math and simulation, and appreciate the power of both
Use a variety of approaches to problem solving
Work with probability concepts algebraically, numerically, and graphically

Rules & Requirements

Prerequisites: Statistics/Computer Science/Information C8, or Statistics/Computer Science C100, or both Stat 20 and Computer Science 61A; and one year of calculus at the level of Mathematics 1A-1B or higher. Corequisite: Mathematics 54, Electrical Engineering 16A, Statistics 89A, Mathematics 110 or equivalent linear algebra

Credit Restrictions: Students will receive no credit for STAT C140 after completing STAT 134. A deficient grade in STAT C140 may be removed by taking STAT 134.

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture, 2 hours of discussion, and 1 hour of supplement per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Final exam required.

Formerly known as: Statistics 140

Also listed as: STAT C140

Probability for Data Science: Read Less [-]

DATA 144 Data Mining and Analytics 3 Units

Terms offered: Fall 2022, Fall 2021, Fall 2020
Data Mining and Analytics introduces students to practical fundamentals of data mining and emerging paradigms of data mining and machine learning with enough theory to aid intuition building. The course is project-oriented, with a project beginning in class every week. The in-class portion of the project is meant to be collaborative and a time for the instructor and GSIs to work closely with project groups to understand the objectives, help work through software logistics, and connect project work to lecture. Lectures will introduce theories, concepts, practical contexts, and algorithms. Students should expect to leave the class with hands-on, contemporary data mining skills they can confidently apply in research and industry.
Data Mining and Analytics: Read More [+]

Objectives & Outcomes

Course Objectives: Conduct manual feature engineering (from domain knowledge) vs. machine induced featurization (representation learning)
Develop intuition in various machine learning classification algorithms (e.g. decision trees, feed-forward neural networks, recurrent neural networks, skip-grams) and clustering techniques (e.g. k-means, spectral)
Foster critical thinking about real-world actionability from analytics
Provide an overview of issues in research and practice that will affect the practice of data science in a variety of domains

Student Learning Outcomes: Develop capabilities in a range of data mining techniques
Gain the ability to solve problems in data mining research and practice
Think critically about how to assess analytics
Use data mining and analytics in a domain of application

Rules & Requirements

Prerequisites: Data 100 (COMPSCI/STAT C100) recommended

Credit Restrictions: Students will receive no credit for DATA 144 after completing INFO 154. A deficient grade in DATA 144 may be removed by taking INFO 154.

Hours & Format

Fall and/or spring: 15 weeks - 3 hours of lecture per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Alternative to final exam.

Instructor: Pardos

Data Mining and Analytics: Read Less [-]

DATA H195A Data Science Honors Thesis Seminar 2 Units

Terms offered: Prior to 2007

The senior honors thesis seminar gives students an opportunity to experience firsthand what it means to do data science research. Over two semesters, students will learn to formulate a research problem, design a research strategy, collect evidence, and write up the findings and analysis. The first semester focuses primarily on the preparation and implementation of a research proposal, as well as data management strategies. During the second semester, we will emphasize analysis and writing. The final result will be a hybrid product with a 20-25 page research paper, with data visualizations and analysis tables, along with a documented data source, annotated code, well documented Github repository, and open science posting of the project.

Data Science Honors Thesis Seminar: Read More [+]

Objectives & Outcomes

Course Objectives: Assist students with project organization and management.
Convey approaches to effective writing and visual communication.
Discuss how to formulate and substantiate an argument with evidence.
Explain approaches to designing a research question and project.
Explore a variety of analytic and visualization techniques and discuss their appropriateness to different research questions.
Identify the challenges in data acquisition and management.

Student Learning Outcomes: Communicate effectively in written, spoken, and graphical form.
Develop an understanding of data availability, constraints, and ethics.
Develop data management skills.
Develop reproducible research and interpret results.
Formulate a proposal for a research project.
Learn how to develop a research question and project.
Understand how to organize empirical work into a written document.
Understand how to use empirical evidence to construct an argument.

Rules & Requirements

Prerequisites: There are no specific prerequisites. Students must be accepted into the data science honors program in order to take this course. Students must complete H195A in order to enroll in H195B

Hours & Format

Fall and/or spring: 15 weeks - 2 hours of seminar per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Alternative to final exam.

Data Science Honors Thesis Seminar: Read Less [-]

DATA H195B Data Science Honors Thesis Seminar 2 Units

Terms offered: Spring 2020

The senior honors thesis seminar gives students an opportunity to experience firsthand what it means to do data science research. Over two semesters, students will learn to formulate a research problem, design a research strategy, collect evidence, and write up the findings and analysis. The first semester focuses primarily on the preparation and implementation of a research proposal, as well as data management strategies. During the second semester, we will emphasize analysis and writing. The final result will be a hybrid product with a 20-25 page research paper, with data visualizations and analysis tables, along with a documented data source, annotated code, well documented Github repository, and open science posting of the project.

Data Science Honors Thesis Seminar: Read More [+]

Objectives & Outcomes

Course Objectives: Assist students with project organization and management.
Convey approaches to effective writing and visual communication.
Discuss how to formulate and substantiate an argument with evidence.
Explain approaches to designing a research question and project.
Explore a variety of analytic and visualization techniques and discuss their appropriateness to different research questions.
Identify the challenges in data acquisition and management.

Student Learning Outcomes: Communicate effectively in written, spoken, and graphical form.
Develop an understanding of data availability, constraints, and ethics.
Develop data management skills.
Develop reproducible research and interpret results.
Formulate a proposal for a research project.
Learn how to develop a research question and project.
Understand how to organize empirical work into a written document.
Understand how to use empirical evidence to construct an argument.

Rules & Requirements

Prerequisites: There are no specific prerequisites. Students must be accepted into the data science honors program in order to take this course. Students must complete H195A in order to enroll in H195B

Hours & Format

Fall and/or spring: 15 weeks - 2 hours of seminar per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Letter grade. Alternative to final exam.

Data Science Honors Thesis Seminar: Read Less [-]

DATA 197 Field Studies in Data Science 1 - 4 Units

Terms offered: Fall 2019
Students take part in organized individual field sponsored programs with off-campus organizations or tutoring/mentoring relevant to specific aspects and applications of data science on or off campus. Note Summer CPT or OPT students: written report required. Course may not count toward major requirements but will be counted in the cumulative units toward graduation.
Field Studies in Data Science: Read More [+]
Rules & Requirements

Prerequisites: Consent of instructor (see department advisor). Upper-division standing

Repeat rules: Course may be repeated for credit with advisor consent.

Hours & Format

Fall and/or spring: 15 weeks - 1-4 hours of fieldwork per week

Summer:
6 weeks - 2.5-10 hours of fieldwork per week
8 weeks - 2-7.5 hours of fieldwork per week
10 weeks - 1.5-6 hours of fieldwork per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Offered for pass/not pass grade only. Alternative to final exam.

Field Studies in Data Science: Read Less [-]

DATA 198 Directed Group Studies for Advanced Undergraduates 1 - 4 Units

Terms offered: Spring 2022, Spring 2021
Written proposal must be approved by a faculty sponsor, who will serve as Instructor of Record. Seminars for the group study of selected topics, which will vary from semester to semester. Topics may be initiated by students.
Directed Group Studies for Advanced Undergraduates: Read More [+]
Rules & Requirements

Prerequisites: Instructors may require students to enroll concurrently or have completed Data 8 (COMPSCI/STAT/INFO C8). Upper-division standing and consent of instructor

Repeat rules: Course may be repeated for credit without restriction. Students may enroll in multiple sections of this course within the same semester.

Hours & Format

Fall and/or spring: 15 weeks - 1-4 hours of directed group study per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Offered for pass/not pass grade only. Final exam not required.

Directed Group Studies for Advanced Undergraduates: Read Less [-]

DATA 199 Supervised Independent Study and Research 1 - 4 Units

Terms offered: Prior to 2007
Independent study and research by arrangement with faculty or staff. This course allows students to obtain course credit for participation in undergraduate research. Students may opt either to participate in a semester-long series of workshops which provide a guided research experience with project milestone assignments and regular feedback, or they may opt to work independently with supervision from one faculty research mentor.
Supervised Independent Study and Research: Read More [+]

Objectives & Outcomes

Student Learning Outcomes: Develop and refine skills acquired in other courses in a hands-on, self-directed research project.
Identify how to properly manage data and describe best practices in programming and analytics.
Integrate feedback from an instructor into research on a regular basis.
Learn how to structure and complete a research project working independently.

Rules & Requirements

Prerequisites: Instructors may require students to enroll concurrently or have completed Data 8 (COMPSCI/STAT/INFO C8). Upper-division standing and consent of instructor

Repeat rules: Course may be repeated for credit without restriction.

Hours & Format

Fall and/or spring: 15 weeks - 3-12 hours of independent study per week

Summer:
6 weeks - 7.5-30 hours of independent study per week
8 weeks - 5.5-22.5 hours of independent study per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Undergraduate

Grading/Final exam status: Offered for pass/not pass grade only. Alternative to final exam.

Supervised Independent Study and Research: Read Less [-]

DATA C200 Principles and Techniques of Data Science 4 Units

Terms offered: Fall 2022, Spring 2022, Fall 2021, Spring 2021, Spring 2020
Explores the data science lifecycle: question formulation, data collection and cleaning, exploratory, analysis, visualization, statistical inference, prediction, and decision-making. Focuses on quantitative critical thinking and key principles and techniques: languages for transforming, querying and analyzing data; algorithms for machine learning methods: regression, classification and clustering; principles of informative visualization; measurement error and prediction; and techniques for scalable data processing. Research term project.
Principles and Techniques of Data Science: Read More [+]

Rules & Requirements

Prerequisites: COMPSCI C8 / INFO C8 / STAT C8 or ENGIN 7; and either COMPSCI 61A or COMPSCI 88. Corequisites: MATH 54 or EECS 16A

Credit Restrictions: Students will receive no credit for DATA C200\COMPSCI C200A\STAT C200C after completing DATA C100.

Hours & Format

Fall and/or spring:
8 weeks - 6 hours of lecture, 2 hours of discussion, and 2 hours of laboratory per week
15 weeks - 3 hours of lecture, 1 hour of discussion, and 1 hour of laboratory per week

Summer: 8 weeks - 6 hours of lecture, 2 hours of discussion, and 2 hours of laboratory per week

Additional Details

Subject/Course Level: Data Science, Undergraduate/Graduate

Grading: Letter grade.

Formerly known as: Statistics C200C/Computer Science C200A

Also listed as: COMPSCI C200A/STAT C200C

Principles and Techniques of Data Science: Read Less [-]