Computational and Data Science and Engineering

The Designated Emphasis (DE) in Computational and Data Science and Engineering Program (CDSE) at the University of California, Berkeley trains students in modeling and high-performance simulation of complex physical systems, as well as in several aspects of data analysis, statistics, machine learning, data visualization, etc. The CDSE program is committed to the development of new curricula and expanded programs aimed at the development and propagation of the use of tools of scientific computation to enhance research across multiple disciplines. To that end, the CDSE program will actively support the training and multidisciplinary education of scientists, engineers, and technical specialists who are experts in relevant areas.

The CDSE program crosses numerous disciplines, and participating departments include all of the departments in the college of engineering, computer science, mathematics, chemistry, astronomy, neuroscience, and political science, among many others.

Upon graduation, the student receives a "PhD in X with a Designated Emphasis in Computational and Data Science and Engineering" on their transcript and diploma. This designation certifies that she or he has successfully completed a designated emphasis in addition to the departmental requirements for the PhD. Completion of the DE-CDSE will also be posted to the student's transcript. We encourage interested readers to visit data.berkeley.edu/decdse (https://data.berkeley.edu/decdse/).

Applicants must already be students within a Ph.D. program at UC Berkeley.

Students must be accepted into the program and petition to add the DE-CDSE before taking their **Qualifying Exam**.

Applicants for the DE-CDSE program are required to:

- Submit an online application; for access to the application, please see the program's website (https://data.berkeley.edu/decdse/).
  - With this form, the student must specify their proposed area of CDSE study and list which three courses (https://data.berkeley.edu/decdse_students/) they will take to fulfill the requirements.
  - Curriculum Vitae (CV)
  - Transcripts (most recent copy of all undergraduate and graduate transcripts)
  - Letter of recommendation from their advisor
  - One-page statement about why they are applying to the program

For further information regarding admission to graduate programs at UC Berkeley, please see the Graduate Division's Admissions website (http://grad.berkeley.edu/admissions/).

Coursework/Curriculum

The student must take one course from each of the categories below; all courses taken to fulfill the DE requirements must be taken for a letter grade.

Category 1: Mathematical Tools

Select one of the following:

MATH 221	Advanced Matrix Computations [4]	
MATH 228A	Numerical Solution of Differential Equations [4]	
MATH 228B	Numerical Solution of Differential Equations [4]	
STAT 204	Probability for Applications [4]	
STAT 210A	Theoretical Statistics [4]	
STAT 210B	Theoretical Statistics [4]	
STAT 215A	Applied Statistics and Machine Learning [4]	
STAT 215B	Statistical Models: Theory and Application [4]	
STAT 232	Experimental Design [4]	
STAT C239A	The Statistics of Causal Inference in the Social Science [4]	
STAT 240	Nonparametric and Robust Methods [4]	
STAT C241A/ COMPSCI C28	Statistical Learning Theory [3]	
STAT C241B/ COMPSCI C281B	Advanced Topics in Learning and Decision Making [3]	
STAT 243	Introduction to Statistical Computing [4]	
STAT 244	Computing for Statistics and Data Science with Julia [2]	
STAT 248	Analysis of Time Series [4]	
A course not listed, by petition to the program director		

Category 2: High Performance Computing

Select one of the following:

COMPSCI C267	Applications of Parallel Computers	3-4
COMPSCI 286A	Course Not Available	4
COMPSCI 289A	Introduction to Machine Learning	4
COMPSCI 294	Special Topics	1-4
CS 294-73 "Software Engineering for Scientific Computing"		
CS 294-127 "Computational Imaging"		
CS 297-143 "Design, Evaluation, and Implementation of Modern Warehouse-Scale Computers"		
STAT 259	Reproducible and Collaborative Statistical Data Science	4
CHEM/BIO ENG C242	Machine Learning, Statistical Models, and Optimization for Molecular Problems	4

Category 3: Application area

Application area courses that utilize the above tools in a significant manner [1]	

[1] The student proposes this course, including a detailed syllabus documenting the use of mathematics and computation in an application area.

Qualifying Exam

Students must have a DE-CDSE component in their qualifying exam, with a DE-CDSE faculty member on the exam committee.

Dissertation

At least one member of the DE-CDSE faculty be on the dissertation committee.