ACADEMIC SENATE
BERKELEY
320 STEPHENS HALL
UNIVERSITY OF CALIFORNIA

March 26, 2018

BENJAMIN HERMALIN
Vice Provost for the Faculty

Subject: Proposal to Establish a Bachelor of Arts Program in Data Science

Dear Ben,

On March 19, 2018, Divisional Council (DIVCO) discussed the proposal cited in the subject line, informed by commentary of the Committee on Budget and Interdepartmental Relations (BIR), and Undergraduate Council (UGC). The committee comments are appended in their entirety for your consideration.

DIVCO endorsed the proposal, however, in doing so, we underscored a number of ongoing concerns.

UGC provided extensive feedback to the proposers which led to the appended clarification. DIVCO shares UGC's concern about ensuring coordination between the Data Science (DS) Division and units offering courses in DS Domain Emphases. We underscore UGC's recommendation:

> We are concerned that the broad array of courses will result in a great deal of back-and-forth between Data Science advisers/leadership and the units offering the courses that make up the Domain Emphases; to the extent this can be accomplished in advance through MOUs, we would have greater confidence in the success of students in the major. At the same time, it is not a good use of resources to establish a large number of formal arrangements before the actual demand for those tracks in the major are understood. It may be advisable to pursue these arrangements with a smaller number of Domain Emphases, however.

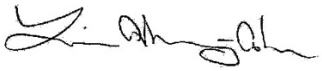
We also agree with UGC's recommendation regarding "truth in advertising" with respect to Domain Emphases that rely on courses that are part of impacted majors or programs. DIVCO recommends that prospective DS majors are explicitly informed

about potential impaction in Domain Emphases. The establishment of MOUs should also help to address this issue.

DIVCO discussed BIR's concern about the need for strategic planning with respect to FTE allocation, noting "it is hard to envision creating new processes and policies while we are still uncertain about the structure of the new Data Science Division." DIVCO agrees. At the same time, we do not want to obstruct progress on establishing the major while the campus resolves these institutional and structural issues. Instead, we request that the new major be formally reviewed by DIVCO at the conclusion of its third year, with particular attention paid to concerns identified in committee reports, including "enrollment, resource availability (advising and TAS), distribution of demand across menus, availability of courses, student satisfaction, faculty engagement, and any other challenges that emerge in the early years."

In sum, DIVCO supports the establishment of a Bachelor of Arts program in Data Science and strongly urges the proposers to be attentive to the identified issues in the committee reports. We look forward to the findings of the three year review.

Sincerely,



Lisa Alvarez-Cohen
Chair, Berkeley Division of the Academic Senate
Fred and Claire Sauer Professor
Department of Civil and Environmental Engineering

Encls. (3)

Cc: Bob Jacobsen, Dean, College of Letters & Science Undergraduate Studies
David Culler, Interim Dean, Data Sciences Division
Michael Lucey, Chair, Committee on Budget and Interdepartmental Relations
Mark Stacey, Chair, Undergraduate Council
Sumei Quiggle, Associate Director staffing Undergraduate Council
Will Lynch, Senate Analyst, Committee on Budget and Interdepartmental Relations
Noam Manor, Institutional Research Analyst, Office of Planning and Analysis

University of California, Berkeley

COMMITTEE ON BUDGET AND
INTERDEPARTMENTAL RELATIONS

February 2, 2018

CHAIR LISA ALVAREZ-COHEN
BERKELEY DIVISION OF THE ACADEMIC SENATE

RE: Proposal to Establish a Bachelor of Arts Program in Data Science

You have asked for our comments on the proposal from the College of Letters & Science (L&S) to establish a Bachelor of Arts degree in Data Science. We are happy to provide our thoughts. Following standard practice, we limit our comments to issues raised by the proposal that are within our purview: that is, to the program's effect on faculty time and its implications for merit advancement, to the program's implications for future allocation of faculty FTE, and to faculty compensation. Silence on other aspects of the proposal should not be read as a commentary on those other aspects.

The proposed Data Science major comprises foundational lower-division courses and a flexible program of upper-division courses, including many possible domain emphases to accommodate students' particular interests and career goals. It is intended to prepare students for a wide range of data analytics professions and for graduate school in related fields. The courses in the proposed major will be staffed by faculty of multiple departments as a Group Major (Berkeley Division of the Academic Senate Regulation 810.B). A proposal for a Bachelor of Science degree in Data Science from the College of Engineering is expected to follow.

We agree with Bob Jacobsen, Dean of the College of Letters & Science Undergraduate Studies, that the proposal describes "a high-quality program that combines depth and breadth while remaining accessible to a wide range of our students." We also agree that there is a clear demand for a Data Science major.

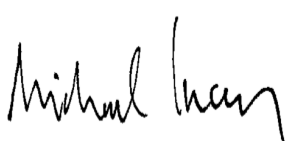
In terms of governance, the proposal states that a "committee of faculty from multiple departments will be formed at the Data Science Divisional level to oversee the major, review the integration of its course offerings, curate its course lists, liaise with other units, programs, and majors, provide a focal point for addressing student and institutional issues, and perform other curricular and administrative tasks." We note that this Division has yet to be constituted, and note as well that it might have been more efficient for both the BA and the BS majors in Data Science to have been evaluated at the same time, and once the Division is in place, or its outlines more clear.

In terms of new FTE, the proposal seems to imply that new faculty in the Departments of Electrical Engineering & Computer Sciences and Statistics will not be required beyond those needed for expansion of Data Science courses for existing, non-Data Science majors. Appendix B of the proposal argues that there will be a need for an additional 3 "teaching faculty,"

presumably Lecturers with Security of Employment (LSOE), over the next 7 years if Data 8 is taken by all students. The Appendix goes on to say, "FTE demand in advanced courses aligns closely with expected regular faculty research needs." The proposal also claims that, "[i]n general, if the Data Science major draws students away from the L&S CS and EECS majors, it will reduce stress on the[se] existing programs." It seems to us more likely that, because of the large number of Data Science courses to be taught by EECS, Statistics, and IEOR faculty, the proposed major will increase stress on all three Departments, at least in the short term, in terms of both teaching and faculty advising of majors. As for other Departments, Appendix B states, "[we] expect [enrollment in Data Science courses] to correlate with data science related FTE requests, which will be based not only on this need, but broader data science aspects of research and teaching in the departments." Appendix B also states, "A critical area of need (recognized in the L&S Executive Committee response) is faculty who can create and teach courses in the Human Context and Ethics category. Minimally two such FTE are needed to carry the additional teaching load for the Data Science major." Whether those FTE would be in the new Division or not is unclear at this point. As David E. Culler, Interim Dean of the Division of Data Sciences, states in his response to the L&S Executive Committee, "We do expect that, as part of building the Division, FTE will be devoted to bringing in faculty to enhance strength in areas that provide HCE courses as part of a confluence of research and teaching. How particular FTE requests that address this goal fall among Divisional requests, those of departments or other academic units within the Division, those of other departments or schools in conjunction with the Division, or joint FTE among these will vary with the particular request, and those processes are currently in development. These FTE requests will go through the conventional Academic Senate processes; in those processes, the establishment of the major and its associated needs will provide additional basis for resource allocation."

We recognize and support the campus's stated priority in Data Science. We also see a critical need for strategic planning across units regarding FTE allocation in this general area, and in the specific area of Human Contexts & Ethics in particular. We regret the absence in the proposal of any indication that discussions with relevant existing units are underway and proving productive. We also concur that this major will be dependent to an as yet unclear degree on temporary academic support (TAS) budgets, and on positions in the LSOE series. We concur with the statement in the report that "[p]rogress on TAS policy will be important for the Data Science major, not just among the departments forming the group, but for a wider set of departments that may see enrollment shifts as they offer or introduce Domain Emphasis and Human Contexts and Ethics courses." We do not see how the major can be viable until these problems are solved. In addition, it is hard to envision creating new processes and policies while we are still uncertain about the structure of the new Data Science Division.

Thank you for the opportunity to comment on this report.



Michael Lucey
Chair

ML/al

ACADEMIC SENATE
BERKELEY
320 STEPHENS HALL
UNIVERSITY OF CALIFORNIA

March 14, 2018

PROFESSOR LISA ALVAREZ-COHEN
Chair, Berkeley Division of the Academic Senate

Re: UGC Comments on L&S Data Science Major Proposal

Dear Chair Alvarez-Cohen,

At our March 7, 2018 meeting, the Undergraduate Council discussed the proposal to create a new Data Science Major in L&S. This discussion was informed by a proposal clarification submitted by Data Science leadership in response to a request from the UGC in February. The proposed major consists of a strong technical foundation in mathematics, computation, and statistics, and is built around a spine of new courses that develop the core material of data science at both the lower and upper division levels. From there, the major becomes considerably less structured, with menu-based requirements in Computational and Inferential Depth (CID), Domain Emphases (DE), and Human Context and Ethics (HCE). The UGC in its discussion of the original proposal voiced concerns over the unconstrained nature of these later requirements in the major. In the requested clarification provided by Data Science leadership, the lists of courses in the Domain Emphases were shortened and the sequencing of these courses was clarified. Further, more details were provided regarding paths through the major and the (hidden) constraints of some prerequisites. Finally, the relation between HCE choice and DE was made explicit.

The UGC is supportive of this major and wants to ensure its success. The clarification helped to define the structure of curriculum, giving us confidence that the curricular choices students can make will result in an intellectually coherent degree that is worthy of UC Berkeley. In particular, the spine of the major – DS8, DS100, DS102, and DS104 – was viewed very positively. The remainder of the degree requirements provide an opportunity for specialization and the pursuit of individual interests, but remain, even after clarification, somewhat unstructured. UGC feels strongly that extensive advising of students in the major will likely be necessary (more below). In the end, the major will train leaders in the field of data science, will almost certainly be well received by employers, and may influence pedagogy well beyond the Berkeley campus.

At the same time, the UGC sees warning signs on the horizon for this major. The supporting materials and justification suggest very extensive student interest, and projections of rapid

growth in the major. The potential for this growth may be offset, to some extent, by the depth and rigor of the technical aspects of the major (making it very similar to CS in many ways), but if the major does grow as projected, the UGC emphasizes the need for ***significant resources for both advising and instruction***. Further, complex interdependence between campus units is fundamental to the successful delivery of this major. Both the deployment of resources and the degree of coordination across campus units will require care and attention during the formative period of the major. To help ensure its success, UGC recommends that there be early check-ins on the major as it develops. In the 3rd year of the major, **UGC would like to formally hear from the Data Science Leadership with a report on enrollment, resource availability (advising and TAS), distribution of demand across menus, availability of courses, student satisfaction, faculty engagement, and any other challenges that emerge in the early years.** We believe that UGC is in a unique position to be able to look across units and majors and to consider the possible interactions or conflicts as they emerge. A mandated and formal review in the 3rd year would coincide with the first set of 4-year majors transitioning into the upper division, so patterns and conflicts will start to emerge.

As part of both preparing the major to launch, and in order to frame the issues that will be important in a review, we wish to provide more information here on the concerns the UGC has raised. Although the clarification was a step in the right direction, the UGC would actually encourage the proposers to go even further in constraining the curriculum in the early years. This suggestion is motivated by trying to ensure: (1) coherence of the students' academic programs; (2) high quality instruction throughout the major and coordination across units; and (3) truth in advertising. We take each of these concerns individually:

Coherence of the academic programs: This concern was addressed to some extent by the clarification, but the UGC believes it may be valuable to go further in constraining the DE options during the early years of the major, but also allowing students to petition for Domain Emphases that are not pre-defined. As these requests emerge, new tracks could be formally identified and mapped out. With the current structure, there will be significant demands on advising resources; investment of these resources must be commensurate with the need. Of course, it may be that through strong advising, or through self-selection, the programs of studies that students pursue will be coherent even with an unconstrained set of requirements. This would be an important point of evaluation in a follow up review.

High quality instruction and ensuring coordination across units: The complexity of the structure of the major in the upper division may create uncertain and uncontrolled demand in upper division courses across many units on campus. Reducing the number of options in the Domain Emphases would allow for more predictable resource allocation, including enrollment slots, classroom space and TAS allocations. We are concerned that the broad array of courses will result in a great deal of back-and-forth between Data Science advisers/leadership and the units offering the courses that make up the Domain Emphases; to the extent this can be accomplished in advance through MOUs, we would have greater confidence in the success of students in the major. At the same time, it is not a good use of resources to establish a large number of formal arrangements before the actual demand for those tracks in the major are understood. It may be advisable to pursue these arrangements with a smaller number of Domain Emphases, however.

Truth in advertising: If Domain Emphases are presented to the students declaring the major, they will expect that the courses listed as part of that track will be available to them. At the moment, it is not clear that this will be possible with the lists provided. Some courses (CompBio175, e.g.) have not been offered in nearly a decade (or more); others are constrained by classroom space/technology and already have waitlists that are comparable to the size of the class itself (Geography 185, e.g.); still others (Business Analytics emphasis) are made up of courses that are part of impacted majors and clear priority enrollment systems are already in place. Without MOUs with the units offering these courses, it is difficult to have confidence that (a) the courses will be offered, and (b) data science majors will have access to them.

Again, the UGC wants to emphasize that we are supportive of this major and our comments and suggestions here are intended to help it to have a successful launch and sustainable future. The major is particularly challenging from the perspective of resources, due to the dependence on many units through the Domain Emphases and Human Context and Ethics requirements. The conflicts that may emerge with students in other majors, or within individual courses, should be anticipated and mitigated in advance to the extent possible.

Sincerely,

Mark Stacey
Chair, Undergraduate Council

Enclosure: Clarification of L&S Data Science Major Proposal

MA/scq

Clarification of L&S Data Science Major Proposal

This document serves be responsive to the concerns in the original proposal (submitted to L&S Executive Committee on September 28, 2017) raised by the Academic Senate Undergraduate Council (UGC) and in doing so to clarify certain specifics of the proposed major. The (unhighlighted) text is extracted from the original proposal, (removing treatment of the numerous issues surrounding the major that were requested for the L&S Executive Committee process and for the OPA process.) Further clarification or modifications to the text of the proposal are highlighted (as here) to better serve the needs of the UGC. In particular, this document sharpens the draft Domain Emphases based on input received in the process and designates certain specifications on the use of HCE courses based on UGC input.

Throughout the clarification of the major the recursive set of prerequisites are explicitly called out, so that none are "hidden". This issue was of leading concern in the faculty-driven process of formulating such an integrated major, discussed at great length in faculty committees, and attended to repeatedly. Further clarification, as suggested, is valuable to students, advisors, departments, and approval committees. Prerequisite demands is complicated by the varying nature of prerequisites within existing programs, but managed here with careful selection of courses. Within each of the building blocks of the program, pathways to completion are present with no prerequisites beyond those explicitly called out in the major, while also retaining pathways with a small number of additional prerequisites that are naturally present in the programs of students with the interest represented by the pathway. Hence the major offers multiple pathways for students with varying degrees of preparation.

Templates of 4, 3, and 2 year programs are "fleshed out" with examples drawing from particular Domain Emphases, showing both that the program and the prerequisites are comfortably satisfied and how selection of Computational and Inferential Depth courses can resonate with Domain Emphasis selection.

All Domain Emphases have been culled and analyzed from the viewpoint of intellectual coherence, prerequisite chains, accessibility, and excellence. Addressing the concerns raised by the UGC has improved both presentation and content of these elements. All DEs identify a lower division (required) entry point and a cohesive list of upper division courses that develop the theme of the particular emphasis with few external requirements.

The principles above guide the formation of the emphases, rather than rigid adherence to a formula, such as exactly one lower division entry point and 3-5 upper division courses. For example, in 19 of the 27 DEs, a single lower division entry is natural and is specified. However, in some, e.g., Ecology and Environment and Geospatial Information and Technology, our current upper division courses draw loosely upon a suite of introductory courses, so picking only one for data science students would be artificial. For some DEs, better alignment will arise as result of the consideration of the data science proposal; for example, certain Sociology courses currently

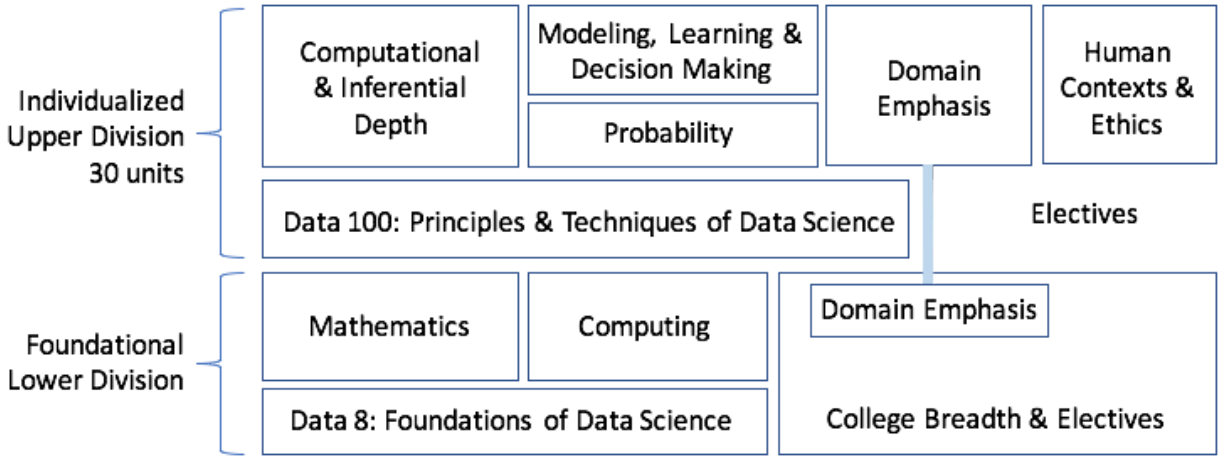
require SOCIOL 1 or SOCIOL 3, but SOCIOL 5 is becoming established and is likely to be more suited for Data Science students. This issue will be resolved through consultation with the department and further curation of the lists, so the more flexible option is retained here.

For 17 of the DEs the upper division list is 4-7 courses; seven have 8 or 9, three have longer lists. Generally, course prerequisites are limited to the entry course and at most one upper division course that is also in the DE, so satisfying the prerequisite completes the DE. In a few cases, a course that is particularly relevant to the domain, say Terrestrial Hydrology in Ecology and Environment, has additional lower division prereqs (e.g., Physics 7A or Chem 1A), so it is included along with alternatives that are free of such requirements. In a few cases, a course with a particular upper division prerequisite outside the DE is included in the list because it would not be reasonable to prevent the student with that additional background the option of utilizing that course in the DE. In the three DEs with longer lists (Quantitative Social Science, Social Welfare, Health, and Poverty, Inequalities in Society) a broad set of courses address either different facets of the issue or common facets from different perspectives. Offering those variations is a distinct benefit to the student. A much greater concern is the two DEs with 4 upper division courses, as the scarcity of offerings can make it difficult for the student to complete the requirement. However, additional appropriate options are not yet available in our course offerings and the emphases are of value to students.

We have adopted the UGC suggestion to divide the current HCE options into a set of four that are available in all DEs, while the remaining three are available selectively in DEs where they are most directly relevant. When created, the new Data 104 course would add another, universal option. It remains to be seen whether benefit to students merits the additional administrative complexity of this restriction. The issue of funneling students into a single HCE offering is of considerable debate. Generally, there has been strong support for drawing on Berkeley's broad and deep intellectual resources on the social, political, economic, and human implications of technology in providing HCE offerings - and concern that there would be "one view" offered to students. The current HCE offerings are taken very seriously by their offering units - and two have been created just this semester by renowned faculty. Prof. Cathryn Carson teaching STS C100 / ISF C100G / HIST C182 Introduction to Science, Technology, and Society: Human Contexts and Ethics of Data and Prof. Deirdre Mulligan is teaching INFO 188 Beyond the Data: Humans and Values. All the HCE courses are in high demand[1] and it remains to be seen whether they can be increased in size while retaining high quality. They certainly can be increased in number and variety. At this stage it is premature - and probably unadvised - to suggest that any one of these, or the additional one in progress, is "the default course" to satisfy this requirement.

[1] A table of past enrollments is provided in the proposal. This term in the first time offerings, 43 students are in Info 88 and 54 in History c182.

Salient Features of the Data Science BA



The major program is designed to provide integrative course experiences in the lower division and upper division, as well as the technical depth in computation and inference required for students to engage in data science upon graduation. The design leverages the wealth of existing courses at Berkeley, while bringing in an essential set of new courses to connect those existing resources in a systematic, intellectually coherent, and logistically feasible manner.

Students are permitted flexibility in choosing an area of emphasis, which may be a domain of data-intensive discovery, a theoretical area, or an integrative intellectual thread. Its flexible structure permits students primarily interested in the systems, mathematics, application, or contextualization of data science to carry out individualized paths of study. The emphases are designed to permit exposure to the ways of thinking of a domain and are not limited to courses that focus on data analytics. It is important for the student to learn to form the bridges between. One of the three courses in the Domain Emphasis is a lower division course; this is typically needed to satisfy prerequisites of the upper division courses in the Emphasis; it will typically satisfy a College or University breadth requirement and starts the entering student considering possible Domains of Emphasis early in their studies.

With the recent introduction of data science courses and the broad array of existing courses that support this integrative major, students could declare and complete the major today, without further course development. However, with the major in place, it will be very desirable to develop at least two additional upper division courses that are particularly tailored to data science students. One of these (Data 102) will be a machine learning course that builds directly on earlier parts of the program and focuses on the theory and application most relevant to modeling and decision making. The other (Data 104) will draw together a range of human and societal reflections most relevant to data science as a broad field of application.

Governance: A committee of faculty from multiple departments will be formed at the Data Science Divisional level to oversee the major, review the integration of its course offerings, curate its course lists, liaise with other units, programs, and majors, provide a focal point for addressing student and institutional issues, and perform other curricular and administrative tasks. In particular, this committee will curate the domain emphases.

Anticipating that existing programs will introduce courses and sequences that would serve well as additional Domain Emphases, a systematic petition process will be provided by which compelling student petitions are generalized to form new, curated lists. This process will involve advising staff and the governance committee. This will avoid students selecting courses that do not form a coherent emphasis. There is considerable precedent for this approach on campus in the highly popular Cognitive Science, Political Economy and new and recently approved Global Studies majors.

Lower Division Requirements (25+ units)

- [4] Foundations of Data Science (CS/Info/Stat C8)
- [8] Calculus (Math 1A/B or equivalent)
- [4] Linear Algebra (Math 54 or EE 16A/16B)
- [2+] Program Structures (CS 88 or CS 61A or E7)
- [4] Data Structures (CS 61B)
- [3+] Domain Emphasis: one course of an area of emphasis will typically be a lower division course (3+ units) and typically will serve to meet a College or University requirement.

Upper Division Requirements (30 units)

- [4] Principles and Techniques of Data Science – Data 100,
- [4] Probability (1 of *Stat 140*, or Stat 134 or EE 126 or IEOR 172)
- [4] Modeling, Learning and Decision-Making (1 of *Data 102*, CS 189, Stat 154, IEOR 142)
- [7+] Computational and Inferential Depth – see description and course list
- [3+] Human Contexts & Ethics – see description and course list
- [7+] Domain Emphasis – see description and course list

The combination of courses must comprise at least 30 units, i.e., at least one additional unit must be taken in one of the latter three categories, and no course may be counted in more than one category[2]

[2] In a few cases a course we find that a course providing computational and inferential depth in a manner broadly valuable to data science students is also an integral part of a domain emphasis. Permitting such flexibility simplifies administration of the major and permits broader student choice.

Preliminary course lists are provided here. They are expected to evolve with time through governance processes as the presence of the major and its constituent courses stimulates other curricular innovations. (Please see the governance section above.) Final determination of course lists may depend on departmental willingness to accommodate data science students. Negotiating such arrangements prior to major approval would be premature. Outside the departments of the group major, departments represented in course lists are generally also interested in having the minor available to their students. Thus, it is reasonable to expect a climate of reciprocity and collaboration, where we work together for the benefit of students generally. Again, this model has been successful at scale in Cognitive Science, Political Economy and Global Studies. It reflects the stronger sense of working together that is central to Berkeley's future.

The typical upper division program, once the two additional courses are developed, will include Data 100, Stat 140, Data 102 for Modeling, Learning, and Decision Making, and Data 104 for Human Contexts and Ethics. However, in place of Data 102, students with particular computer science interests may elect CS 189 (with CS 70 and Math 53). Similarly, those with a Statistics interest may elect STAT 154 (with STAT 135), or with an Operations Research bent, IEOR 142 (with IEOR 165 or an equivalent course in Statistics). The earliest cohorts will need to pursue these options. (Approximately 900 students a year are taking one of these courses recently.) Recognizing that some pathways through the major may require an additional course to meet prerequisites is another reason the major is kept to the minimum allowable for the Group Major, 30 upper division units. The forthcoming Human Contexts and Ethics course is not expected to introduce any issues in meeting prerequisites, and none of the existing HCE courses have any prerequisites.

Summary of prerequisites (from guide.berkeley.edu) for current courses serving the central pillar of the major.

- Data 100 (CS/STAT C100): CS/INFO/STAT C8 or ENG 7; and either CS 61A or CS 88. Corequisite: MATH 54 or EE 16A
- STAT 140: CS/INFO/STAT C8 and one year of calculus at the level of Math 1A-1B or higher. Co-requisite: Math 54, EE 16A, or equivalent linear algebra course
- STAT 134: One year of calculus
- EE 126: CS 70 preferred but not required. Familiarity with linear algebra
- IEOR 172: Students should have a solid knowledge of calculus, including multiple variable integration, such as Math 1A-1B or 16A-16B, as well as programming experience in Matlab or Python
- CS 189: Math 53 and 54; CS 70 or consent of instructor
- STAT 154: Math 53 and 54 or equivalents; STAT 135 or equivalent; experience with some programming language. Math 55 or equivalent exposure to counting arguments is recommended but not required. (This may be changed to allow just STAT 140.)
- IEOR 142: IEOR 165 or equivalent course in statistics. Prior exposure to optimization is helpful but not strictly necessary. Some programming experience/literacy is expected

Grandfathering: To better support current students interested in declaring the major, for 5 semesters following approval of the major the Data 8 requirement will be viewed as fulfilled by either of the following course pairs: Stat 135 + CS 61A, Stat 20 or 21 + CS 61A. The Data 100 requirement is viewed as fulfilled by the topics course CS194-16 Introduction to Data Science or by STAT 133 + STAT 135 + CS 61A . The inclusion of existing courses as options to fulfill specific major categories handles the remaining issues of phasing in the new major. Thus, assuming the major is approved Spring 2018, these measures will terminate Spring 2020.

Computational and Inferential Depth

A student will be required to take two courses comprising 7 or more units from a list of advanced courses providing computational and inferential depth (C&ID) beyond that provided in Data 100 and the lower division. It is recognized that, currently, some of these courses have prerequisites that are not formally within major, so for some combinations a student may need to use electives to complete those. However, many options are available that do not place such demands. And efforts will be made (with the offering departments) to re-examine the prerequisites of the CID courses to increase the number of paths that do not require such additional courses.

The following provides the initial list of C&ID courses. The list will be curated by the governance committee and updated as course offerings evolve. Where present, prerequisites beyond those required in the major are shown in brackets, with recursive expansion. The choice of C&ID courses will influence how a student elects to fulfill lower division requirements, in particular whether Math 54 or EE 16A/B is used for linear algebra. Advising staff will need to understand and convey these dependencies. These will be documented clearly and communicated to students early on to ensure a feasible and timely path to degree.

For several of the C&ID courses, prerequisites are expected to be streamlined as the major matures. Importantly, Statistics has already adjusted its upper division prerequisites to so that STAT 140 can be utilized in place of 134 (that change has not yet propagated to guide.berkeley.edu). With some additional course development, this process can continue as the major comes fully into being. In particular, data science students will have the probability portion of CS 70 (Discrete Mathematics and Probability Theory) from other courses, needing only the discrete math portion. Similarly, since CS 61C is a required part of the CS and EECS majors, it is frequently a prerequisite for upper division courses, although some of those courses depend on only a portion of the material covered in CS 61C. And likewise, most upper division statistics courses assume STAT 134 and 135, but these are expected to be streamlined to allow greater access for data science majors. The governance committee for the Data Science major will continue to work with the relevant departments to help make these changes possible.

Course List: Computational and Inferential Depth (C&ID)

- CS 161. Computer Security (4 units) [CS 61C, CS70 or Math 55]
- CS 162. Operating Systems and Systems Programming (4 units) [CS 61C, CS 70]

- CS 164. Programming Languages and Compilers (4 units) [CS 61C]
- CS 168. Introduction to the Internet: Architecture and Protocols (4 units) [CS 162]
- CS 169. Software Engineering (4 units)) [CS 61C, CS 70 or Math 113]
- CS 170. Efficient Algorithms and Intractable Problems (4 units) [CS 70]
- CS 186. Introduction to Database Systems (4 units) [CS 61C]
- CS 188. Introduction to Artificial Intelligence (4 units) [CS 70]
- EECS 127. Optimization Models in Engineering (4 units) [EE 16A/B or consent]
- EE 120. Signals and Systems (4 units) [EE 16A/B]
- EE 123. Digital Signal Processing (4 units) [EE 120, EE 16A/B]
- EE 129. Neural and Nonlinear Information Processing (4 units) [EE 120, EE 16A/B]
- ESPM 174 Design and Analysis of Ecological Research (4 units)
- IEOR 115 Industrial and Commercial Data Systems (3 units)
- IEOR 135 Applied Data Science with Venture Applications (3 units)
- INFO 159 Natural Language Processing (3 units) [CS70 or Math 55]
- INFO 190-1 Introduction to Data Visualization (3 units)
- NUC ENG 175 Methods of Risk Analysis (3 units)
- PHYSICS 188. Data science and Bayesian statistics (4 units) (pending COCI approval)
- STAT 135. Concepts of Statistics (4 units)
- STAT 151A. Linear Modelling: Theory and Applications (4 units) [STAT 135]
- STAT 152. Sampling Surveys (4 units)
- STAT 153. Introduction to Time Series (4 units)
- STAT 158. The Design and Analysis of Experiments (4 units) [STAT 135 concurrently]
- STAT 159. Reproducible and Collaborative Statistical Data Science (4 units) [Currently STAT 133, 135, to be revised]

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. The purpose of this requirement is to equip the student with an understanding of the human and social structures, formations, and practices that shape data science activity (such as data collection and analysis, data stewardship and governance, work to ensure privacy and security, deployment of data in societal or organizational settings, decision-making with data, engagements of data with justice, practices of data ethics) and to allow them to gain experience and practice with modes of critical thinking, reflection, and engagement with these experiences and the choices involved.

While the set of options available to Berkeley students to fulfill this requirement are varied, there is a need for additional seats to meet the likely demand. There are also sets of shared issues that show up in diverse contexts around the ethics of data, algorithmic development, acquisition, communication, and decision-making. As the major develops, an additional Data 104 course will be developed by faculty to address these aspects. This course will be developed by faculty with expertise in this area to address generalizable as well as domain-specific issues, in a way

that will lay a broad foundation for students across careers that may move between domains. It is expected that once Data 104 is developed, it will likely become the most common way of satisfying the HCE requirement. However, it will remain valuable to students to have other options available to them, allowing them to chose to draw from the wealth of relevant perspectives on these topics offered within Berkeley.

In order to be included on the list of courses that satisfy the HCE requirement, a course syllabus needs to give significant attention to data analytics (as at least one among several foci); to provide students access to structured forms of academic inquiry in the humanities, social sciences, or related professional fields; and to engage them in some form of reflective inquiry, writing, analysis, project work, or practice that surfaces questions of individual or societal choices and supports making reasoned ethical choices in complex situations. The governance of the major will include a process for reviewing class syllabi for inclusion on the HCE list.

The four courses on the main HCE course list below can be utilized with any domain emphasis. The three additional HCE courses can be used in conjunction with certain Domain Emphases, with exceptions by petition. It is hoped that as other departments expand their offerings in data science methods, they will adapt or develop courses to meet HCE criteria in relation to a Domain Emphasis in their area.

None of these HCE courses have prerequisites.

Course List: Human Contexts and Ethics

- AMERSTD/AFRICAM C134. Information Technology and Society (4 units)
- INFO 188. Beyond the Data: Humans and Values (3 units)
- ISF 100J. The Social Life of Computing (4 units)
- STS C100 /ISF C100G / HIST C182. Introduction to Science, Technology, and Society: Human Contexts and Ethics of Data (4 units)

For certain Domain Emphases, additional HCE options are available:

- BIO ENG 100. Ethics in Science and Engineering (3 units)
- CY PLAN 101. Introduction to Urban Data Analytics (4 units)
- ESPM C167/PUB HLTH C160. Environmental Health and Development - 4 units

Domain Emphases

Domain Emphases are integral to the Data Science major. A Domain Emphasis may be a domain of data-intensive inquiry, a theoretical area, or an integrative intellectual thread. The student is permitted substantial flexibility in choosing an emphasis and courses within it. Each domain of emphasis will be defined by a list of courses, as described below. These courses need not come from a single department. We expect that faculty in various domains of data-intensive inquiry, theoretical foundations of data science, or its contextualization will contribute to the major by defining additional emphasis lists. This integrative aspect of the major

is intrinsic the field and naturally brings many parts of the university into relationship with Data Science.

An emphasis will typically be a focused topic; it is roughly half of a minor. A well-formed 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.

A domain emphasis will typically draw from four or more courses so that it is resilient and does not require coordinated course scheduling, i.e., the student should have ample opportunity to complete two upper division courses. The more frequently offered, the shorter the domain list needs to be and vice versa. The governance of the major will include a process for introduction of new Domain Emphases and periodic review of Domain Emphasis lists.

The DEs shown below have been further curated since the original proposal, and prerequisites for each course are explicitly stated here. Those DEs that most naturally accommodate additional HCE courses are designated by the line

HCE add: <courses>.

Courses that are currently available only to students in an established major have been removed. No DE course has an upper division prerequisite chain within the DE longer than one course.[3]

Domain emphases are rooted in a lower division course, which is typically also a prerequisite for the advanced courses. *The lower division course (**shown in bold**) is a required element of the DE.*

For most DE's a single such course is specified. In cases where current naturally related upper division courses state a non-specific or one of several lower division entries, we have allowed the student to choose one from a short list, e.g., Ecology and Environment.

Computational Biology Methods

- **BIOLOGY 1A. General Biology (4 units)**
- CMPBIO 175. Introduction to Computational Biology and Precision Medicine (3 units) [ ]
- MATH 127. Mathematical and Computational Methods in Molecular Biology (4 units) [Math 53, 55]

[3] Although many such advanced courses relevant to DEs exist and would be of interest to data science students, students would have fulfilled the DE in meeting the prerequisites for such a course, so it would be an elective and there is no need to include it in the DE course list.

- COMPSCI 176. Algorithms for Computational Biology (4 units) [CS 70, CS170]
- BIO ENG 131. Introduction to Computational Molecular and Cell Biology (4 units) [Bio 1A]
- BIO ENG 143. Computational Methods in Biology (4 units) [Math 53]
- BIO ENG 144. Introduction to Protein Informatics (4 units) [ ]

HCE add: BIO ENG 100, ESPM C167/PUB HLTH C160

Computational Molecular Biology

- **BIOLOGY 1A. General Biology (4 units)**[4]
- MCELLBI 102 Survey of the Principles of Biochemistry and Molecular Biology (4 Units) [BIO 1A]
- MCELLBI 104. Genetics, Genomics and Cell Biology (4 units) [MCELLBI 102]
- MCELLBI 132. Biology of Human Cancer (4 units) [MCELLBI 102]
- MCELLBI 140. General Genetics (4 units) [BIO 1A]
- MCELLBI/PLANTBI C148. Microbial Genetics and Genomics (4 units) [MCELLBI 102]
- MCELLBI 143. Evolution of Genomes, Cells and Development (3 units) [MCELLBI 102]
- MCELLBI 149. The Human Genome (3 units) [MCELLBI 104 or 140]

HCE add: BIO ENG 100, ESPM C167/PUB HLTH C160

Human Biology

- **BIOLOGY 1A. General Biology (4 units)**
- MCELLBI 50. Immune System and Disease (4 units) [remove, chem]
- INTEGBI 131. General Human Anatomy (3 units) [Bio 1A, Bio 1B or Chem 1A]
- INTEGBI 132. Survey of Human Physiology (4 units) or MCELLBI 136. Physiology (4 units) [INTEGBI 131]
- INTEGBI 164. Human Genetics and Genomics (4 units) [Bio 1A, 1B]
- MCELLBI 160. Cellular and Molecular Neurobiology (4 units) [1A, co-req Physics 8AB]
- MCELLBI 165. Neurobiology of Disease (3 units) [MCELBI 160]

HCE add: BIO ENG 100, ESPM C167/PUB HLTH C160

Population Health and Environment

- **BIOLOGY 1A. General Biology (4 units)**
- DEMOG 110. Introduction to Population Analysis (3 units) [ ]
- ESPM C103 / INTEGBIO C156. Principles of Conservation Biology (4 units) [Bio 1A]
- PUB HLTH 150A. Introduction to Epidemiology and Human Disease (4 units)
- ESPM 108B. Environmental Change Genetics (3 units) [Bio 1A]

---

[4] Please note, all DEs start with **the** required lower division course, **shown in bold**. Upper division courses have numbers 100 or larger. Where multiple lower division entry points are options, it is called out explicitly.

- ESPM 111. Ecosystem Ecology (4 units) [Bio 1A]
- INTEGBI 153. Ecology (3 units) [Bio 1A]
- INTEGBI 170LF. Methods in Population and Community Ecology (3 units) [INTEGBI 153]

HCE add: BIO ENG 100, ESPM C167/PUB HLTH C160, CY PLAN 101

Toxicology and Disease

- **BIOLOGY 1A. General Biology (4 units)**
- NUSCTX 110. Toxicology (4 units) [Bio 1A]
- NUSCTX 121. Computational Toxicology (3 units) [Bio 1A]
- NUSCTX 160. Metabolic Bases of Human Health and Diseases (4 units) [Bio 1A]
- PUB HLTH 150A. Introduction to Epidemiology and Human Disease (4 units)

HCE add: BIO ENG 100, ESPM C167/PUB HLTH C160

Ecology and Environment

One of the following lower division courses[5]

- **ESPM 2. The Biosphere (3 units)**
- **ESPM 6. Environmental Biology (3 units)**
- **ESPM 15. Introduction to Environmental Sciences (3 units)**
- **GEOG 40. Introduction to Earth System Science (4 units)**
- **ESPM/L&S C46. Climate Change and the Future of California (4 units)**
- **ESPM 88B. Data Sciences in Ecology and the Environment (2 units)**

Upper division

- ESPM 100. Environmental Problem Solving (4 units) [ecology]
- ESPM 102B. Natural Resource Sampling (2 units) [ ]
- ESPM 102BL. Laboratory in Natural Resource Sampling (2 units) [ ]
- ESPM C104 / ENVECON C115. Modeling and Management of Biological Resources (4 units) [ ]
- EPS/ESPM C129. Biometeorology (3 units) [ ]
- GEOG C136/ESPM C130. Terrestrial Hydrology (4 units) [Chem 1A, Physics 7A, or consent]
- ESPM 157. Data Science in Global Change Ecology (3 units) [Data 8]
- ESPM C170 / EPS C183. Carbon Cycle Dynamics (3 units) [ ]
- ESPM/EPS C180 / CIV ENG C106. Air Pollution (3 units) [Chem 1AB, Physics 7A, or consent]

HCE add: BIO ENG 100, ESPM C167/PUB HLTH C160, CY PLAN 101

[5] Please note, this DE is an example of where the student is best served by multiple lower division entry points, of which one is required.

Evolution and Biodiversity

- **BIOLOGY 1A (4 units) and/or 1B (4 units)** depending on UD choices[6]
- ESPM/INTEGBI C105. Natural History Museums and Biodiversity Science (3 units) [ ]
- ESPM 108B. Environmental Change Genetics (3 units) [Bio 1AB]
- ESPM 152. Global Change Biology (3 units) [Bio]
- INTEGBI 113L. Paleobiological Perspectives on Ecology and Evolution (4 units) [ ]
- INTEGBI 141. Human Genetics (4 units) **OR** INTEGBI 164. Human Genetics and Genomics (4 units)
- INTEGBI 160. Evolution (4 units) [1B] **OR** INTEGBI 167. Evolution and Earth History: From Genes to Fossils (4 units) [1AB]
- INTEGBI 161. Population and Evolutionary Genetics (4 units) [1AB]
- INTEGBI 162. Ecological Genetics (4 units) [1B]
- INTEGBI 172. Coevolution: From Genes to Ecosystems (4 units) [1A or 1B]

HCE add: BIO ENG 100, ESPM C167/PUB HLTH C160

Environmental Economics

- **ENVECON C1 / ECON C3. Introduction to Environmental Economics and Policy (4 units)**
- ENVECON 100. Microeconomic Theory with Application to Natural Resources (4 units) [C1]
- ENVECON C101 / ECON C125. Environmental Economics (4 units) [ENVECON 100]
- ENVECON C115 / ESPM C104. Modeling and Management of Biological Resources (4 units) [ ]
- ENVECON 141. Agricultural and Environmental Policy (4 units) [ENVECON 100]
- ENVECON 142. Industrial Organization with Applications to Agriculture and Natural Resources (4 units) [ENVECON 100]
- ENVECON 145. Health and Environmental Economic Policy (4 units) [ENVECON 100]
- ENVECON 147. Regulation of Energy and the Environment (4 units) [ENVECON 100]

HCE add: ESPM C167/PUB HLTH C160, CY PLAN 101

Geospatial Information and Technology

One of the following lower division courses

- **ESPM 72. Introduction to Geographic Information Systems (3 units)**
- **GEOG 80. Digital Worlds: An Introduction to Geospatial Technologies (4 units)**
- **GEOG 88. Data Science Applications in Geography (2 units)**
- **ESPM 88A. Exploring Geospatial Data (2 units) (likely to evolve to 3 units)**

---

[6] In this DE, natural upper division courses as for one of the other biology entry. A student will often choose to take both out of interest. Sufficient paths are possible with either.

- **CIVENG 88. Data Science for Smart Cities (2 units)**
- **EPS 50 The Planet Earth 4 Units as required per upper division**

Upper division

- GEOG 183. Cartographic Representation (5 units) [ ]
- GEOG 185. Earth System Remote Sensing (3 units) [ ]
- GEOG 187. Geographic Information Analysis (4 units) [ ]
- GEOG/LD ARCH C188. Geographic Information Systems (4 Units) [ ]
- ESPM 137. Landscape Ecology (3 units) [ ]
- ESPM 164. Geographic Information Systems for Environmental Science and Management (3 units) [ ]
- ESPM 172. Photogrammetry and Remote Sensing (3 units) [ ]
- ESPM 173. Introduction to Ecological Data Analysis (3 units) [ ]
- EPS 101. Field Geology and Digital Mapping (4 units) [eps 50]

HCE add: CY PLAN 101

Physical Science Analytics

- **PHYS 77. Introduction to Computational Techniques in Physics (3 units)**[7]

- ASTRON 121 Radio Astronomy Laboratory (4 units) [Math 53, Phys 7ABC]
- EPS 109. Computer Simulations in Earth and Planetary Sciences (4 units) [ ]
- EPS 122. Physics of the Earth and Planetary Interiors (3 units) [Phys 7AB, Math 53]
- EPS C183 / ESPM C170. Carbon Cycle Dynamics (3 units) [ ]
- GEOG C136 / ESPM C130. Terrestrial Hydrology (4 units) [Phys 7A, Chem 1A]
- GEOG C139 / EPS C181. Atmospheric Physics and Dynamics (3 units) [Math 53, Phys 7ABC]
- NUC ENG 101 Nuclear Reactions and Radiation (4 Units) [Phys 7C]
- NUC ENG 130. Analytical Methods for Non-proliferation (4 units) [NUC ENG 101]
- NUC ENG 155. Introduction to Numerical Simulations in Radiation Transport (3 units) [Math 53]

HCE add: BIO ENG 100

Digital Humanities and Data Arts

One of lower division

- **ART W23AC. Data Arts (4 units)**
- **HISTORY 88. How Does History Count (2 units) (likely to evolve to 3 units)**
- **L&S 88. Rediscovering Texts as Data (2 units) (to be regularized within a**

[7] Additional Physics prerequisites are required for the entry and upper division courses. Those are unavoidable and it is unlikely that a student would consider this DE without having taken the lower division physics sequence.

departmental rubric)

Upper Division

- HISTORY 100. Text Analysis for Digital Humanists and Social Scientists (4 units) (piloted as a topics course, to be regularized)
- HISTORY 100. Calculating Americans: Big Histories of Small Data (4 units) (piloted as a topics course, to be regularized)
- HISTORY 104. The Craft of History (4 units)
- HISTART 192DH Digital Imaging and Forensic Art History (4 units)
- THEATER 166 / NWMEDIA 190. Making Sense of Cultural Data (units TBD) (piloted as a topics course, to be regularized)
- ENGLISH/HISTART C181. Digital Humanities, Visual Cultures (4 units)
- NESTUD 190. Introduction to Digital Humanities: From Analog to Digital (units TBD) (piloted as a topics course, to be regularized)

Econometrics[8]

- **ECON 1. Introduction to Economics (4 units)**
- ECON 100A. Economic Analysis--Micro (4 units) [ECON 1]
- ECON/MATH C103. Introduction to Mathematical Economics (4 units) [MATH 53]
- ENVECON/IAS C118. Introductory Applied Econometrics (4 units) [ ]
- ECON 119. Psychology and Economics (4 units) [ECON 100A]
- ECON 121. Industrial Organization and Public Policy (4 units) [ECON 100A]

HCE add: CY PLAN 101

Business Analytics

- **ECON 1. Introduction to Economics (4 units)**
- UGBA 101A Microeconomic Analysis for Business Decisions (3 Units) [ECON 1]
- UGBA 104. Analytic Decision Modeling Using Spreadsheets (3 units)
- UGBA 103 Introduction to Finance (4 Units)
- UGBA 137. Introduction to Financial Engineering (3 units) (piloted as a topics course, to be regularized) [UBA 103]
- UGBA 141. Production and Operations Management (3 units) [UBA 104]
- UGBA 106 Marketing 3 Units [ ]
- UGBA 161. Marketing Research: Data and Analytics (3 units) [UBA 106]
- UGBA 190T. Topics in Business Analytics and Data Mining (3 units) (to be regularized)
- UGBA 192L. Applied Impact Evaluation (2 units)

HCE add: CY PLAN 101

---

[8] Further advanced courses are natural fits, but have long prereq chains. ECON 140 Economic Statistics and Econometrics requires ECON 100AB or 101AB and ECON/PUB POL C142 / POL SCI C131A. Applied Econometrics and Public Policy (4 units) requires ECON 140.

Quantitative Social Science

One of lower division as needed[9]

- **SOCIOL 1 Introduction to Sociology (4 Units)**
- **SOCIOL 3AC Principles of Sociology: American Cultures (4 Units)**
- **SOCIOL 5. Evaluation of Evidence (4 units), or**
- **POLI SCI 3. Introduction to Empirical Analysis and Quantitative Methods (4 units)**

Upper Division

- ECON C110. Game Theory in the Social Sciences (4 units)
- ENVECON C118/IAS C118. Introductory Applied Econometrics (4 units)
- DEMOG 110. Introduction to Population Analysis (3 units)
- SOCIOL 106. Quantitative Sociological Methods (4 units) [SOCIO 5]
- MASSC 130. Research Methods in Media Studies (4 units) [MEDIAST 10]
- PHILOS 141. Philosophy and Game Theory (4 units)
- POLI SCI 133. Selected Topics in Quantitative Methods (4 units)
- DEMOG/SOCIOL C126. Sex, Death, and Data (4 units) [SOCIOL 1 or 3]
- SOCIOL 165. Social Networks (4 units) [SOCIOL 1 or 3]
- DEMOG/ECON C175. Economic Demography (4 units) [ECON 1]
- DEMOG 180. Social Networks (3 units)

HCE add: CY PLAN 101

Computational Linguistics and Natural Language Processing

- **LINGUIS 5 Language and Linguistics (4 Units)**
- LINGUIS 110. Introduction to Phonetics and Phonology (4 units) [LINGUIS 100]
- LINGUIS 113. Experimental Phonetics (3 units) [LINGUIS 110]
- LINGUIS C160 / COGSCI C140. Quantitative Methods in Linguistics (4 units) [LINGUIS 100]
- INFO 159. Natural Language Processing (3 units) [CS70 or MATH 55]

Psychology and Cognition

- **PSYCH 1. General Psychology (3 units)**
- PSYCH 110 Introduction to Biological Psychology (3 Units) [PSYCH 1]
- COG SCI C100 Basic Issues in Cognition (3 Units)
- PSYCH 127. Cognitive Neuroscience (3 units) [MCELLBI C61 or PSYCH 110 or COG SCI C100]
- PSYCH 134. Health Psychology (3 units) [PSYCH 1 and MCELLBI C62]
- PSYCH 143. Language Acquisition (3 units) [ ]
- PSYCH 167. Stigma and Prejudice (3 units) [PSYCH 1]

[9] The lower division entry of this DE reflects the current state of transition in lower division sociology offerings. This issue is expected to be resolved by the time the major becomes available for students.

Organizations and the Economy

- **SOCIOL 1 Introduction to Sociology (4 units), or**
- **SOCIOL 3AC Principles of Sociology: American Cultures (4 units)**
- SOCIOL 110. Organizations and Social Institutions (4 units) [Soc 1 or 3]
- SOCIOL 120. Economy and Society (4 units) [Soc 1 or 3]
- SOCIOL 116. Sociology of Work (4 units) [Soc 1 or 3]
- SOCIOL 119S. Organizational Strategy and Design (4 units) [Soc 1 or 3]
- SOCIOL 121. Innovation and Entrepreneurship (4 units) [Soc 1 or 3]
- GWS 139. Women, Gender, and Work (4 units) [ ]

Inequalities in Society

- **SOCIOL 1 Introduction to Sociology (4 units), or**
- **SOCIOL 3AC Principles of Sociology: American Cultures (4 units)**
- AFRICAM 101. Research Methods for African American Studies (4 units) [intro stat]
- AFRICAM 111. Race, Class, and Gender in the United States (3 units) [ ]
- ETH STD 101A. Social Science Methods in Ethnic Studies (4 units) [ ]
- ETH STD 101B. Humanities Methods in Ethnic Studies (4 units) [ ]
- GWS 131. Gender and Science (4 units) [ ]
- GEOG C155 / AFRICAM C156. Race, Space, and Inequality (4 units) [ ]
- SOCIOL 130 / SOCIOL 130AC. Social Inequalities (4 units) [SOCIOL 1 or 3]
- PSYCH 167. Stigma and Prejudice (3 units) [PSYCH 1]
- PUB POL C103. Wealth and Poverty (4 units) [ ]
- PUB POL 117AC. Race, Ethnicity, and Public Policy (4 units) [ ]
- SOCIOL 113. Sociology of Education (4 units) [SOCIOL 1 or 3]
- SOCIOL 124. Sociology of Poverty (4 units) [Intro Soc]
- SOCIOL 131. Race and Ethnic Relations (4 units) [SOCIOL 1 or 3]
- SOCIOL 133. Sociology of Gender (4 units) [SOCIOL 1 or 3]

HCE add: CY PLAN 101

Social Policy and Law

- **SOCIOL 5. Evaluation of Evidence (4 units)**
- PUB POL 101. Introduction to Public Policy Analysis (4 units)
- GWS 132AC. Gender, Race, and Law (4 units)
- POL SCI 124M. The Scientific Study of International Conflict (4 units)
- SOC WEL 112. Social Welfare Policy (3 units)
- SOC WEL 181. Social Science and Crime Prevention Policy (3 Units)
- SOCIOL 111P. Families, Inequality and Social Policy (4 units) [SOCIOL 1 or 3]
- LEGALST 102. Policing and Society (4 units)
- LEGALST 160. Punishment, Culture, and Society (4 units)
- POLECON 111. Poverty and Social Policy (3 units)

HCE add: CY PLAN 101

Social Welfare, Health, and Poverty

- **SOCIOL 5. Evaluation of Evidence (4 units)**
- GPP 105. The Ethics, Methods, and Pragmatics of Global Practice (4 units) [ ]
- GPP 115. Global Poverty: Challenges and Hopes in the New Millennium (4 units) [ ]
- GLOBAL 102. Critical Thinking in Global Studies (4 units) [ ]
- GWS 130AC. Gender, Race, Nation, and Health (4 units) [ ]
- LEGALST 158. Law and Development (4 units) [ ]
- PB HLTH 112. Global Health: A Multidisciplinary Examination (4 units) [ ]
- PB HLTH/SOCIOL C155. Sociology of Health and Medicine (4 units) [SOCIOL 1 or 3]
- PB HLTH C160 / ESPM C167. Environmental Health and Development (4 units) [ ]
- PB HLTH 181. Poverty and Population (3 units) [ ]
- POLECON 111. Poverty and Social Policy (3 units) [ ]
- SOCIOL 115G. Global Health and Social Justice (4 units) [SOCIOL 1 or 3]
- SOC WEL 112. Social Welfare Policy (3 units) [ ]
- INFO 181. Technology and Poverty (3 units) [ ]

HCE add: CY PLAN 101

Urban Planning and Sustainable Development & Engineering

- **CIV ENG 11. Engineered Systems and Sustainability (3 units), or**
- **LD ARCH 12. Environmental Science for Sustainable Development (4 units), or**
- **CIVENG 88A. Data Science for Smart Cities (2 units)**
- CIV ENG 107. Climate Change Mitigation (3 units) [ ]
- CIV ENG 110. Water Systems of the Future (3 units) [ ]
- CIV ENG 156. Infrastructure Planning and Management (3 units) [CIVENG 93]
- LD ARCH 130. Sustainable Cities and Landscapes (4 units) [ ]
- GEOG 181. Urban Field Study (4 units) [ ]
- GEOG/LD ARCH C188. Geographic Information Systems (4 units) [ ]
- CY PLAN 119. Planning for Sustainability (3 units) [ ]

HCE add: CY PLAN 101

Applied Mathematics and Modeling

- **MATH 53. Multivariable Calculus (4 units)**
- MATH 110. Linear Algebra (4 units)
- MATH 128A. Numerical Analysis (4 units) [MATH 53]
- MATH 128B. Numerical Analysis (4 units) [MATH 110, 128A]
- IEOR 162. Linear Programming and Network Flows (3 units) [MATH 53, MATH 55]
- CS 267. Applications of Parallel Computers (3 units) [ ]

Industrial Analytics

- **MATH 53. Multivariable Calculus (4 units)**
- ENGIN 120. Principles of Engineering Economics (3 units) [ ]
- IEOR 115. Industrial and Commercial Data Systems (3 units) [ ]
- IEOR 130. Methods of Manufacturing Improvement (3 units) [Probability]
- IEOR 166. Decision Analytics (3 units) [Probability]
- IEOR 173. Introduction to Stochastic Processes (3 units) [Probability]
- UGBA 190T. Topics in Business Analytics and Data Mining (3 units) (to be regularized)
- NUC ENG 167. Risk-Informed Design for Advanced Nuclear Systems (3 units) [ ]

Optimization and Efficiency

- **MATH 53. Multivariable Calculus (4 units)**
- EECS 127. Optimization Models in Engineering (4 units) [EE 16AB]
- EE 144. Fundamental Algorithms for Systems Modeling, Analysis, and Optimization (4 units) [CS 70]
- IEOR 160. Nonlinear and Discrete Optimization (3 units) [MATH 53]
- IEOR 162. Linear Programming and Network Flows (3 units) [MATH 53 and 55]
- IEOR 130. Methods of Manufacturing Improvement (3 units)
- IEOR 190. Introduction to Optimization Modeling (units TBD) (to be regularized as IEOR 164)

Robotics

- **MATH 53. Multivariable Calculus (4 units)**
- CS 188. Introduction to Artificial Intelligence (4 units) [CS70]
- EE 120. Signals and Systems (4 units) [EE 16AB]
- EE C106A / BIO ENG C125. Introduction to Robotics (4 units) [EE120]
- EE C106B / BIO ENG C125B. Robotic Manipulation and Interaction (4 units) EE C106A]
- EE 129. Neural and Nonlinear Information Processing (3 units) [EE 120]
- EE C128 / MEC ENG C134. Feedback Control Systems (4 units) [EE 120]

Cognition and Artificial Intelligence

- **COG SCI 1. Introduction to Cognitive Science (4 units) or**
- **MCELLBI/PSYCH C61. Brain, Mind, and Behavior (3 units) or**
- **MCELLBI/PSYCH C64. Exploring the Brain: Introduction to Neuroscience (3 units)**

- COMPSCI 188. Introduction to Artificial Intelligence (4 Units) [CS 70]
- COG SCI C100 / PSYCH C120 Basic Issues in Cognition (3 units) [ ]
- COG SCI/PSYCH C127 Cognitive Neuroscience (3 units) [COGSCI C100]
- COG SCI 131. Computational Models of Cognition (4 units) [COGSCI 1, CS 70 or Math 55]

- COG SCI 190. Data Science and Cognition (3 units) (piloted as a topics course, to be regularized) [ ]
- PSYCH 114 Biology of Learning (3 units) [ ]
- PSYCH 122 Introduction to Human Learning and Memory (3 units) [ ]

Computational Imaging

- **EL ENG 16B Designing Information Devices and Systems II (4 Units)**
- COMPSCI 184. Foundations of Computer Graphics (4 units) [ ]
- EECS C145B / BIOE C165. Medical Imaging Signals and Systems (4 units) [EE 16AB]
- NUC ENG 107. Introduction to Imaging (3 units) [ ]
- EPS 122. Physics of the Earth and Planetary Interiors (3 units) [PHYSICS 7AB, MATH 53]
- BIO ENG 164. Optics and Microscopy (4 units) [PHYSICS 8AB]
- HISTART 192DH. Digital Imaging and Forensic Art History (4 units) [ ]

Course lists for the data science major are likely to evolve, permitting both richer and more streamlined programs of study, as additional faculty and departmental programs become more connected to the program. For example, course offerings in visualization and in deep application to scientific investigation are currently somewhat limited, but new offerings are in development.

Sample Programs

The following illustrate sample programs under various student entry conditions and goals. Each of these admit plenty of whitespace for rearrangement of schedule and insertion of electives. They are not suggested scheduling, but demonstrate clear feasibility.

Four Year, no Advanced Placement

F FA	Data 8	R&C B	Math 1A	
F SP		American Cultures	Math 1B	CS 61A
S FA		L&S Breadth/Domain Emphasis		CS 61B
S SP		L&S Breadth	Math 54	
J FA	Data 100	L&S Breadth	Probability	

J SP	Data 102	L&S Breadth	Computational & Inferential Depth	
S FA	Human Contexts & Ethics	Domain Emphasis		
S SP		Domain Emphasis	Computational & Inferential Depth	

The following shows a concrete version of this schedule for student with a Domain Emphasis in Population Health and Environment. She decides to take CS 170 and STAT 152 for C&I depth. This (currently) causes her to take an additional lower division course, CS 70, but she has ample room for electives.

The longest of the C&ID prerequisites would add CS 61C to this schedule somewhere before the final semester. Any student considering an L&S CS major would have taken those two courses as part of the lower division. The schedule has plenty of flexibility, so any of the DEs could have been chosen as the concrete example.

F FA	Data 8	R&C B	Math 1A	CS 88
F SP		American Cultures	Math 1B	
S FA		DE Bio 1A		CS 61B
S SP	CS 70	L&S Breadth	Math 54	
J FA	Data 100	L&S Breadth	Probability	
J SP	Data 102 or CS 189	L&S Breadth	C&ID STAT 152	
S FA	HCE Hist C182C/STS C100	DE DEMOG 110		
S SP		DE INTEGBI 153	C&ID CS 170	

A second concrete version of this schedule can be displayed for student with a Domain Emphasis in Environmental Economics. This student decides to develop optimization and modeling more fully, so they elect EE 16AB to complete the linear algebra aspects and follow up with Stat 135 and EE 127 (Optimization Models in Engineering) to complete their computational and inferential depth.

F FA	Data 8	R&C B	Math 1A	
F SP		American Cultures	Math 1B	CS 88
S FA		DE ENVECON C1 / ECON C3	EE 16A	CS 61B
S SP		L&S Breadth	EE 16B	
J FA	Data 100	L&S Breadth	Probability	
J SP	Data 102	L&S Breadth	C&ID EE127	
S FA		DE ENVECON 100	C&ID STAT 135	
S SP	HCE Data 104	DE ENVECON C101 or C115 or 141		

In a third concrete version of this schedule, for student with a Domain Emphasis in Computational Imaging. The student takes on the broader Physics lower division.

F FA	Data 8	R&C B	Math 1B	Physics 7A
F SP	CS 88	American Cultures	EE 16A	
S FA	CS 61B	DE EE 16B		Physics 7B
S SP		L&S Breadth		Math 53
J FA	Data 100	L&S Breadth	Probability	
J SP	Data 102	L&S Breadth	C&ID EE120	
S FA		DE EPS 122	C&ID EE123	
S SP	HCE Data 104	DE BIOE 164[10]		

[10] Physics 7A/B replacing 8A/B.

Three Year, with AP Calculus

F FA	Data 8	R&C B	Math 1B	CS 61A
F SP		American Cultures	Math 54	CS 61B
S FA	Probability	L&S Breadth/Domain Emphasis	Domain Emphasis	
S SP	Data 100	L&S Breadth	Domain Emphasis	
J FA	Human Contexts & Ethics	L&S Breadth	Computational & Inferential Depth	
J SP	Data 102	L&S Breadth	Computational & Inferential Depth	

As a concrete version of this tighter schedule, consider a student with a Domain Emphasis in Organizations and the Economy. They select C&ID courses, ESPM 174 Design and Analysis of Ecological Research, and INFO 190-1 Introduction to Data Visualization (3 units). Electives are still available for any of the DEs or the additional prerequisites of the longer C&ID chains. The extreme case would be the use of EE 123 or 129, because of the need for EE 120, but such a student graduating in 3 years is also taking more than 15-16 units in some semester.

F FA	Data 8	DE SOCIOL 3AC	Math 1B	CS 61A or CS88
F SP			Math 54	CS 61B
S FA	Probability: Stat 140	L&S Breadth/Domain Emphasis	DE SOCIOL 120	
S SP	Data 100	L&S Breadth	DE SOCIOL 116	
J FA	HCE INFO 188	L&S Breadth	C&ID ESPM 174	
J SP	Data 102	L&S Breadth	C&ID INFO 190-1	

Junior Transfer

J FA	Data 8	American Cultures	Math 54	CS 61A
J SP	Human Contexts & Ethics	L&S Breadth/Domain Emphasis	Probability	CS 61B
S FA	Data 100	Domain Emphasis	Computational & Inferential Depth	
S SP	Data 102	Domain Emphasis	Computational & Inferential Depth	

A junior transfer into a data science major can be expect to have taken the equivalent of Calculus 1A, 1B and probably CS 61B Data Structures. The schedule shown here assumes only the former. Increasingly, linear algebra is available in the community colleges as well, but that is not assumed. They select a Domain Emphasis in Inequalities in Society.

J FA	Data 8	DE SOCIOL 3AC	Math 54	CS 88
J SP	HCE Data 104 or Hist C182C/STS C100	DE PUB POL 117AC	Probability	CS 61B
S FA	Data 100		C&ID INFO 159	
S SP	Data 102	DE AFRICAM C156	C&ID IEOR 135	