Expanding Potential 2015-6 Seed Project Awardee Final Report
The purpose of this final report is for readers to understand the motivations, processes, and outcomes for the seed project awardee's programming in order to sustain and scale these efforts.

Please outline your project. When responding, take into consideration if someone wanted to reproduce your program is reading this answer.

The project was first developed by the Energy and Resources Group (ERG) Student Diversity Committee as a way to address the 'leaky pipeline' of underrepresented groups and quantitative skills in STEM fields. The committee saw a disparity in terms of experience and fluency with quantitative analysis by race, gender, and class, and wanted to encourage inclusive education that would also bring attention to these issues.

Over the spring and summer of 2015, this group of graduate students sought partners on campus and designed a curriculum that would eventually come to be called 'Data and Diversity'. In the fall of 2015, we offered the project-based course to undergraduates, with the UC Berkeley Division of Equity and Inclusion and the Berkeley Institute for Data Science (BIDS) as our project clients. After an initial phase of data skills-building and discussion sections on diversity issues, the course culminated in two projects selected by the students with input from the clients. These projects examined the effect of STEM "weeder" courses on enrollment of underrepresented minorities, and the factors leading female PhD students to give up pursuit of an academic career. The students presented their work at the Expanding Potential conference in January 2016.

The ERG Student Diversity Committee intends to offer the course again in fall 2016 while also finding ways to institutionalize the course so that it can be a permanent offering for undergraduates. In addition, the committee is seeking ways to maximize impact of the course, potentially by convening diversity, equity, and inclusion offices from Bay Area campuses to discuss diversity data availability and what communities can do with this data.

What were your initial goals for your program?

Our initial goal was to use data science to better understand how STEM departments can create more equitable, inclusive, and diverse learning environments. We wanted to build the programming skills of underrepresented undergraduates, and to promote campus-wide awareness of diversity statistics in STEM fields through the final projects in the course.

We had further sub-goals for those participating in the course:

For the course clients:

- Leverage data, institutional knowledge, and students' personal experiences to expand knowledge and understanding of disparate experiences of students in STEM fields.
- Identify strategies to advance inclusion, equity, and diversity in STEM (or in a particular department, i.e. BIDS) at Berkeley, assess these strategies using existing data or data collecting during the project.
- Begin a project that will be extended beyond this course of mutual benefit to the client and the students.

For the students:

- Introduce issues of equity, inclusion, and diversity in academics (and society as a whole) and contribute to improving our understanding of these issues at UC Berkeley specifically.
- Introduce basic data analysis and computer programming.
- Provide a structure in which students follow a project through from initial ideation to construction of final deliverables.

How have your goals changed? What motivated these changes?

Our original goal of teaching data science became less emphasized over the course of the semester. The focus of the course shifted towards the diversity data projects because of how much our students were interested in them. From the end-of-semester feedback the students gave us, students generally felt that readings and discussions were covered well, but the programming sessions could have been better planned. In the future, we intend to teach the data section better without sacrificing any of the perspective gained from exploring diversity issues in our discussions.

How does your program address diversity and inclusion issues? How is your program enacting change?

There is a critical need for better understanding of diversity statistics in STEM fields in order to identify and design policies to promote underrepresented groups in STEM. By working with motivated clients across UC Berkeley's campus, we intended for the results of our analysis to be useful for informing campus actions regarding admission, retention, and graduation of students of different socio-economic backgrounds. We presented our findings to key decision-makers, and our students were able to showcase their work at the Expanding Potential conference. At the same time, we aimed to increase the data fluency of Berkeley undergraduates who may not otherwise have had a chance to do hands-on data analysis. These skills are crucial for today's job market, and we believe that early peer-to-peer learning in a supportive, diverse environment builds skills and confidence. By analyzing data on diversity in our own campus community, we also enabled undergraduates to become skilled and informed advocates for diversity in STEM. We are continuing to seek opportunities to continue to highlight the students' work, and we plan to teach the course again in Fall 2016.

What partnerships were necessary to successfully launch this project?

Our clients were essential to the success of the course. The summer before the course was offered, we reached out to numerous groups across campus to find a client with whom collaboration would provide mutual benefit. We eventually partnered with the Berkeley Institute for Data Science, which provided an additional source of skills training and institutional support for our students, and the Division of Equity and Inclusion, which helped us understand data availability and access on campus, and guided us towards interesting and relevant research questions.

How can your program be scaled and applied to different institutions?

Our curriculum is publicly available on our website (datadiversity.berkeley.edu), and could easily be adopted at different institutions. We communicated the failures and successes of our class to a wide audience at the Expanding Potential conference, highlighting the most effective ways to design data analysis projects for diversity data. We also hope the future iterations of the course

can encourage students to go beyond the UC Berkeley campus, perhaps as ‘diversity data consultants’, to help an even wider range of institutions examine their own diversity data.

How did the seed project funding help your project?

Seed project funding was essential for class materials used for brainstorming and visualization, for recruiting skilled data analysis mentors to help us teach the class, for providing snacks at each class (which was held in the evening), and compensating the students for printing their final poster. We intend to use remaining funds to offer the course again, continue to help students communicate their results, and convene a diversity data workshop that would encourage other Bay Area campuses to analyze and communicate their diversity statistics in a way that can enable positive change in their communities. We continue to seek other funding opportunities and collaborators to scale up the project.

Describe how feedback from the Expanding Potential Workshop helped inform your project.

The Expanding Potential conference and workshop was most helpful in forging connections with like-minded experts. Many of them gave us concrete tips, resources, and advice for future iterations of the course, or expressed interest in taking or helping to teach the course. In addition, the design thinking workshop that we ran at the conference helped us see how our research design methods would work for other people from other institutions. We were encouraged that our idea could in fact galvanize others and scale up to different institutions.

What advice would you give to someone initiating this project at another institution?

Having an institutional client who is willing to work with you is essential; otherwise you may face lack of interest or even pushback against results. Impactful projects are surprisingly easy to find. Designing a rigorous data science curriculum takes time. So does recruiting motivated students.

Please summarize any additional results that are not mentioned in the above answers.

N/A

Quote/testimonial

“The Expanding Potential program not only provided support for the Data and Diversity class throughout the semester, but also gave our students a platform to communicate their results to a diverse, engaged audience at the annual conference.”

“The students and clients involved in the Data and Diversity projects were all able to produce powerful, compelling results from simple analysis of Cal’s diversity data - in just a few short weeks, they were able to highlight some major diversity-related problems on campus that had previously received little attention.”

From a student in the class: “The readings and the discussions really opened up my perspectives a lot, and it really enabled me to think about taking courses similar to these more in the future, because not everything is about data science; it's also about the society we're applying to, and what changes we want to make. So honestly while I joined the class to learn more about data science methods, I think I gained the most in discussions, readings, and perspectives.”