

Inclusive Excellence Activities

Undergraduate research experiences empower critical thinking, build the pipeline of future scholars, and open students to thinking of themselves as scholars. These experiences open up the inner workings of the academy and allow undergraduate students to see the pathway from undergraduate education to graduate education. Yet studies show that only a small proportion of students from historically underrepresented backgrounds participate in undergraduate research, even though they benefit the most from these experiences.¹ Since joining SFSU, I have worked to make undergraduate research experiences, especially those that require data analytic skills, more accessible to a broader group of students.

My faculty position at SFSU was part of an interdisciplinary cluster hire with faculty in the Biology department to foster collaboration across the Economics and Biology departments in the use of ‘Big Data’ and Machine Learning methods. My cluster co-hire, Dr. Rori Rohlf, and I realized that an area of shared interest was helping increase the number of undergraduate students who could participate in our respective research labs. Yet we faced a critical common barrier: very few students had the necessary coding skills to engage in the research projects we were leading. We wanted to make data science accessible and ensure that a new cadre of students, especially students from historically underrepresented groups, could fully take advantage of undergraduate research experiences.

We partnered with senior colleagues, Drs. Márquez-Magaña & Akom to write, secure and implement a training grant from NIH/NIMHD focused on encouraging historically underrepresented students to use data and analytic tools to conduct health-relevant research.² With the grant resources, we organized a nine-week summer internship for students to learn to code in R and Python.³ We placed students into semi-autonomous groups of 4-5 undergraduates and paired them with a paid near-peer mentor with more experience. The student teams spent six weeks learning the basics of R/Python using curated online resources. After the initial six weeks, the faculty spent three weeks guiding students in simple research assignments. Each team was assigned a small research project that required students to organize several files and create a visualization using the skills they acquired from the online modules. Students were required to present their results at our summer research symposium.

To ensure the program was inclusive, we used evidence-based practices to encourage students from historically underrepresented groups and female students to apply. We emphasized in our advertisements that we welcomed applicants with no prior coding experience, there was no grade point average (GPA) or prerequisite requirements, and applicants did not need letters of recommendation. We made the application process short and straightforward, asking more questions about the student’s motivations and goals than prior skills. In the program advertisements, we used visuals showcasing diverse students to signal that all students can belong to the program.⁴

An aim of the program was to match students who wanted to find research positions with faculty mentors. We invited faculty from several departments during the program to meet with students through regular lunches. These lunches exposed students to various data science fields— biology, economics, public health, sociology, and ethnic studies. At the end of each summer, we advertised the program to faculty, highlighting that the program participants had now acquired 90 hours of coding/research experience and

¹ For a review report see: Finley, A., & McNair, T. (2013). [Assessing underserved students’ engagement in high-impact practices](#).

² NIH Grant: 1R25MD011714 HEART & SOUL: Enabling full representation in biomedical Big Data science. Role Co-I.

³ Our Biology colleague, Dr. Pennings, had started a version of the program in 2015 with just 5 students. Dr. Rolf and I worked with Dr. Pennings to expand the program to include 30-50 students and to expand the disciplinary lens. I participated from 2016 to 2019. Due to COVID and changes in funding, the program was limited to biology and chemistry students in summer 2020.

⁴ For details of the program see: Pennings P, Banuelos MM*, Catalan FL*, Caudill VR*, Chakalov B*, Hernandez S*, Jones J*, Okorie C*, **Modrek S**, Adelstien N, Rofls R. Ten Simple Rules for an Inclusive Summer Coding Program for Non-Computer-Science undergraduates. *PLOS Computational Biology* 2020, 16(9): e1007833.

were looking to continue to gain further experience. We organized a networking and “speed-matching” research lunch for faculty and program participants to find one another.

Over the three years I participated in the program, we trained over 90 students, most of whom were from historically underrepresented groups. Notably, most program alumni could find paid campus research positions by the end of the program. Over 15 alumni have published papers with their faculty mentors; many more have since gone to graduate school.