SE&IS White Paper on the UCLA DataX Initiative

Gregory Leazer & Ozan Jaquette

July 7, 2020

Vision

text

Plan

Instructional programs

Research and outreach

Increasing faculty capacity and developing new academic programs will help SEIS contribute to DataX research goals. Additionally, because SEIS is unique amongst UCLA academic units in its engagement with public and non-profit organizations across Los Angeles, DataX can improve the ability of SEIS to benefit our local community.

We briefly highlight a handful of centers and projects as illustrative of work that engages with and will benefit from DataX.

- **CRESST**, a research center within SEIS, [ADD ONE SENTENCE ON WHAT CRESST IS]. ADD 1-2 SENTENCES ON HOW CRESST CONTRIBUTES TO AND BENEFITS FROM DATAX.
- **Data for Democracy**. Center X (located within SEIS) created the Data for Democracy project as part of the UCLA Centennial Initiative. The project enables Los Angeles K-12 students to collect and analyze data on critical issues such as housing inequality and access to parks. K-12 students write policy briefs on these topics, published in English and Spanish.
- **Introduction to Data Science (IDS)**. Developed in 2013 as a partnership between Los Angels Unified School District, Center X, and the UCLA Department of Statistics, Introduction to Data Science (IDS) is a one-year course that teaches that teaches computational and statistical thinking for secondary students, with a particular emphasis on reaching students underrepresented in STEM. The IDS data analysis labs teach K-12 students how to manage and analyze data using R.

DataX will increase the analytic capabilities of researchers, graduate students, and K-12 students working on the *Data for Democracy* and *Introduction to Data Science* projects.

More broadly, developing the data science methodological coursework benefit SEIS students and the social justice oriented research and outreach projects they contribute to. For example, several SEIS graduate students who completed the two-course "fundamentals of programming using R" sequence have GSR positions analyzing data from the UCLA Community School. After completing the courses, these students fundamentally improved data quality and efficiency by writing R scripts that create analysis variables used across many analyses. Additionally, the informal Advanced R summer reading group is an offshoot of the nascent R programming sequence. To practice the programming skills they were learning, the reading group prepared data and graphs for a group of CU-Boulder alumni seeking to put pressure on the university to enroll more African American students.

Appendices

Appendix A: Curriculum

AX: Programming and data science using R sequence

text

Appendix B: Research and Outreach

BX: xxxxx

text