

Alexander Franks¹, Ambuj Singh¹, Michael Ludkovski¹, Sang-Yun Oh¹, Sharon Solis¹, Trevor Ruiz¹, Tim Robinson¹, and Nathalie Guebels² ¹ UC Santa Barbara, ² Santa Barbara City College



CAL POLY



Overview

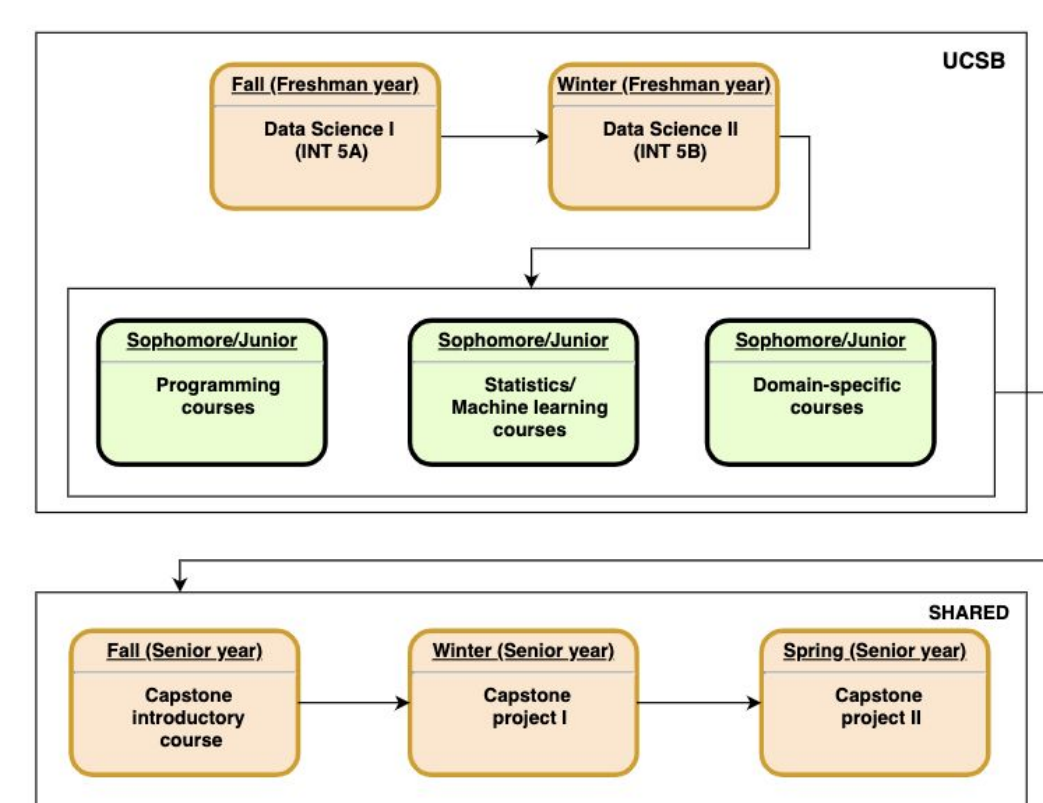
Our project is focused on data science training connecting three main public higher education institutions in California: UC, CSU, and CA Community Colleges. Specifically we,

- developed close collaboration and connections between multiple institutions
- provided stipends as a vehicle for UG training
- focused on developing a stack of best practices:
 - Intro Course development
 - Capstone organization
 - Seminars
 - Assessment
- Collaboration between UCSB and data science fellowship at Cal Poly SLO
 - Cal Poly ran a similar program
 - Joint capstone poster showcase in May 2022
 - Hackathon



New Data Science Courses

- Intro to Data Science** at UCSB
 - Inspired by UC Berkeley's Data 8, but split over 2 quarters
 - CS 5A: focus on Python fundamentals, data visualization, and hypothesis testing
 - CS 5B: introduction to bootstrapping, confidence intervals, regression and classification techniques
 - GenEd credits (pending approval) and related pre-major credits in relevant domain areas
 - Enrollment: 651 since Fall 2020
- Intermediate Data Science** at UCSB (PSTAT 100)
 - DS lifecycle, including statistical inference and prediction, and decision-making
 - Enrollment: 269 since Spring 2021
- Data Science for All** at SBCC (CS 118 / Math 118)
 - New in Fall 2021
 - Approved as GE and transferable to UC and CSU
- Data Science Seminar** at Cal Poly (DATA 472)



Data Science Fellows and Projects

UCSB DS Fellows

- 66 DS Fellows over three years
- 12 URM and 10 Hispanic
- Predominantly juniors and seniors
- Role as "ambassadors of data science"
- \$5k stipend, funded by NSF
- Participate in Capstone Projects
- Fellows participated in "committees"
 - Outreach*: DS presentations to local high schools, community colleges
 - Education*: help w/ intro DS courses
 - Events*: organize seminars, information sessions, poster showcases
 - Infrastructure*: help develop and explore tools and infrastructure for DS classes and research

Capstone Course

- Meet weekly for at least 1 hour with the project sponsor and DS mentor
- In-class time:
 - Discussions about project workflows, communication, teamwork, ethics, etc.
 - Regular updates via blog posts + in class presentations
- Approximately 60 students per year in the three-quarter sequence

Summer Research Experience

- UCSB Summer Immersive Data Science Fellowship for SBCC+CSUSB students
- 4 students each in 2021 and 2022
- Students receive stipends for 8-week internships
- Seminars on professional skills development
 - Abstract and poster creation
 - Lightning talks
 - Research symposium



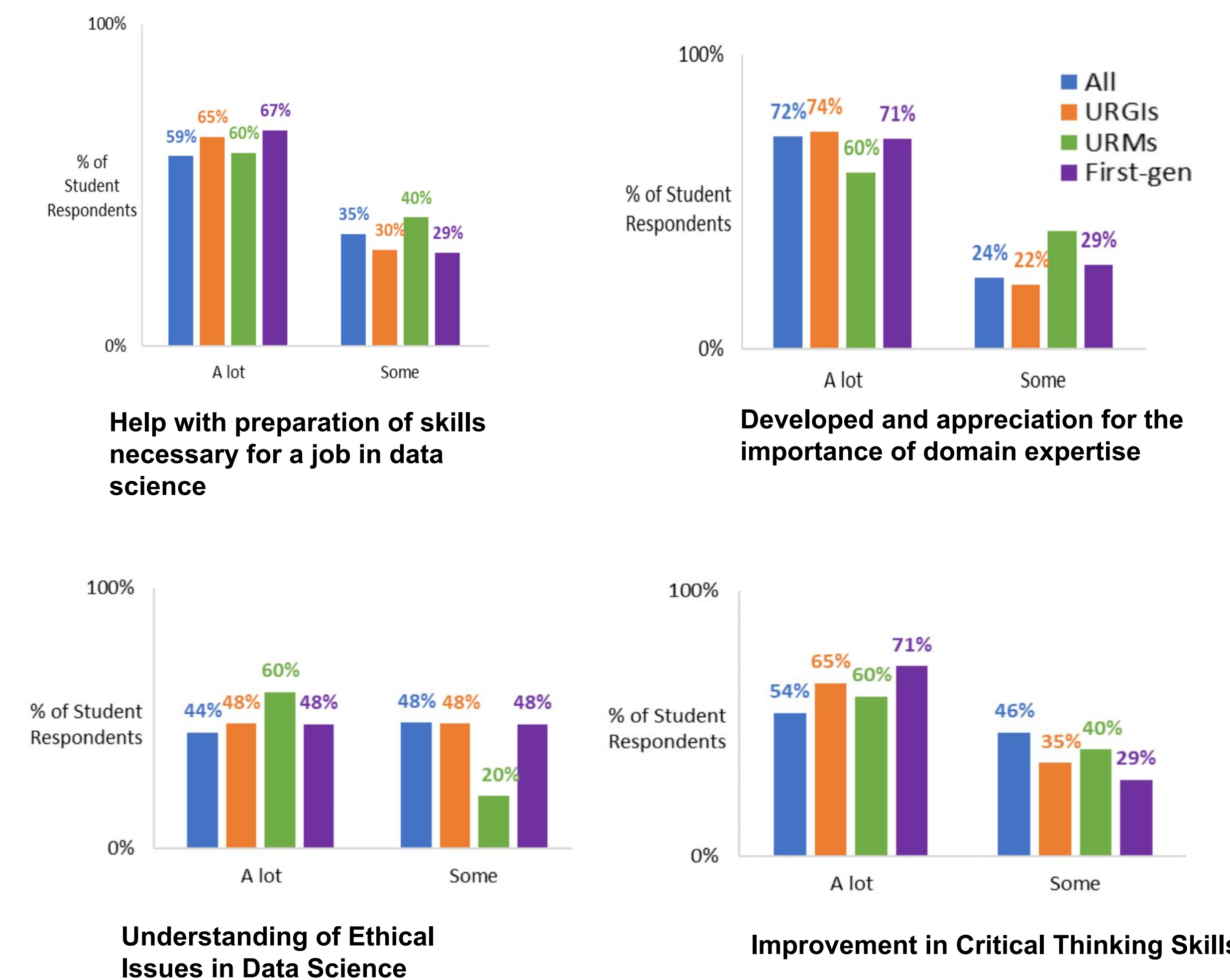
Capstone Projects (2022)

Quantifying Security - Visualizing Usability of Allthenticate	Stress indicators for frontline COVID-19 healthcare workers
Demographic Predictors of Climate Opinion	Product recommendation on scraped usage data
Modeling Neurodegeneration with Single Cell RNA Sequencing	Deep Learning Recommendation System for Property Management Software
Combining computer vision and satellite imagery to map center pivot irrigation	Compression of NLP-Domain Deep Learning Models
Deep image reconstruction from human brain activity	Wearables Predict Stress in Frontline Healthcare Workers
Multi-class claims activity classification based on HTML data	Big Bee: Hair Recognition & Quantification
Impacts of a documentary: The Social Dilemma - analyzing changes in social media usage behavior	Interactive spatio-temporal visualization of long-term physical characteristics of California coastal waters

Assessment

Quarterly meetings with internal assessor allows us to track the impact of our DS classes across campus.

Self assessments of fellows completing the capstone sequence (2022):



Blue = All (n=60), Orange = URGI (n=26), Green = URM (n=6), Purple = First-Generation (n=22). "Very Little" and "None" responses are omitted from the visualization (response percentages generally low single digits)..

Lessons Learned

- Students found immense value in participating in capstone projects.
- In-person allowed for more cohesion.
- Need more emphasis on early training to engage more URM and non-STAT/CS majors.
- Need for collaborative learning assessment, peer-evaluation, and individual accountability in team setting.
- Cal Poly: "Do fewer things well." Recognize that fellowship is extracurricular.
- Stepping stone to larger S-STEM lead by UCSB: PALiSaDS: Pacific Alliance for Low-Income Inclusion in Statistics and Data Science (paliisads.org).



More about our project!